



Intuitive Analog Circuit Design



**CD-ROM
INCLUDED**

Marc T. Thompson



Intuitive Analog Circuit Design

Intuitive Analog Circuit Design

*A Problem-Solving Approach
using Design Case Studies*

By

Marc T. Thompson, Ph.D.



ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Newnes is an imprint of Elsevier



Newnes

Newnes is an imprint of Elsevier
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
Linacre House, Jordan Hill, Oxford OX2 8DP, UK

Copyright © 2006, Elsevier Inc. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone: (+44) 1865 843830, fax: (+44) 1865 853333, e-mail: permissions@elsevier.com.uk. You may also complete your request on-line via the Elsevier homepage (<http://elsevier.com>), by selecting "Customer Support" and then "Obtaining Permissions."



Recognizing the importance of preserving what has been written,
Elsevier prints its books on acid-free paper whenever possible.

Library of Congress Cataloging-in-Publication Data

Thompson, Marc T.

Intuitive analog circuit design.

p. cm.

ISBN-13: 978-0-7506-7786-8 (pbk. : alk. paper)

ISBN-10: 0-7506-7786-4 (pbk. : alk. paper) 1. Electronic circuit design.

I. Title.

TK7867.T48165 2006

621.3815--dc22

2005036750

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

ISBN-13: 978-0-7506-7786-8

ISBN-10: 0-7506-7786-4

For information on all Newnes publications
visit our Web site at www.books.elsevier.com

06 07 08 09 10 10 9 8 7 6 5 4 3 2 1

Printed in the United States of America

Working together to grow
libraries in developing countries

www.elsevier.com | www.bookaid.org | www.sabre.org

ELSEVIER

BOOK AID
International

Sabre Foundation

In memoriam

To my mother, Minnie Ann Thompson,
who bought me my first transistors when I was a child.

Dedication

To Lisa and Sophie M., for your love and patience.
The book is finally done. Mazel tov!

Contents

Preface	<i>xi</i>
Chapter 1: Introduction and Motivation	1
The Need for Analog Designers.....	1
Some Early History of Technological Advances in Analog Integrated Circuits	2
Digital vs. Analog Implementation: Designer's Choice	6
So, Why Do We Become Analog Designers?	8
Note on Nomenclature in this Text	9
Note on Coverage in this Book	9
References	10
U.S. Patents	12
Chapter 2: Review of Signal-Processing Basics.....	13
Review of Laplace Transforms, Transfer Functions and Pole-Zero Plots	13
First-Order System Response	15
Second-Order Systems.....	23
Review of Resonant Electrical Circuits	33
Use of Energy Methods to Analyze Undamped Resonant Circuits	34
Transfer Functions, Pole/Zero Plots and Bode Plots	36
Risetime for Cascaded Systems	37
Chapter 2 Problems	38
References	42
Chapter 3: Review of Diode Physics and the Ideal (and Later, Nonideal) Diode	43
Current Flow in Insulators, Good Conductors and Semiconductors	43
Electrons and Holes	45
Drift, Diffusion, Recombination and Generation.....	48
Effects of Semiconductor Doping	53
PN Junction Under Thermal Equilibrium	54
PN Junction Under Applied Forward Bias.....	57

Contents

Reverse Biased Diode	61
Ideal Diode Equation	62
Charge Storage in Diodes	63
Charge Storage in the Diode Under Forward Bias	65
Reverse Recovery in Bipolar Diodes	65
Reverse Breakdown	66
Taking a Look at a Diode Datasheet	67
Some Quick Comments on Schottky Diodes	70
Chapter 3 Problems	71
References	74
Chapter 4: Bipolar Transistor Models	75
A Little Bit of History	75
Basic NPN Transistor	76
Transistor Models in Different Operating Regions	79
Low-Frequency Incremental Bipolar Transistor Model	82
High-Frequency Incremental Model	85
Reading a Transistor Datasheet	88
Limitations of Hybrid-Pi Model	93
Chapter 4 Problems	94
References	95
Chapter 5: Basic Bipolar Transistor Amplifiers and Biasing	97
The Issue of Transistor Biasing	97
Some Transistor Amplifiers	101
Chapter 5 Problems	121
References	125
Chapter 6: Bandwidth Estimation Techniques and the Method of Open-Circuit Time Constants	127
Introduction to Open-Circuit Time Constants	127
Transistor Amplifier Examples	132
Chapter 6 Problems	156
References	159
Chapter 7: Advanced Transistor Amplifier Techniques	161
Worst-Case Open-Circuit Time Constant Calculations	161
High-Frequency Output and Input Impedance of Emitter-Follower Buffers	168
Bootstrapping	177
Short-Circuit Time Constants	184

Pole Splitting.....	196
Chapter 7 Problems.....	202
References.....	205
Chapter 8: High-Gain Bipolar Amplifiers and BJT Current Mirrors	207
The Need to Augment the Hybrid-Pi Model.....	207
Base-Width Modulation.....	209
Finding Parameters from a Transistor Datasheet.....	211
Common-Emitter Amplifier with Current Source Load	212
Building Blocks	213
Chapter 8 Problems.....	238
References.....	240
Chapter 9: Introduction to MOSFET Devices and	
 Basic MOS Amplifiers	241
Some Early History of Field-Effect Transistors.....	241
Qualitative Discussion of Basic MOS Devices.....	242
Figuring Out the V/I Curve of a MOS Device	243
MOS Small-Signal Model (Low Frequency).....	246
MOS Small-Signal Model (High Frequency)	247
Basic MOS Amplifiers	248
Chapter 9 Problems.....	266
References.....	268
Chapter 10: Bipolar Transistor Switching and the	
 Charge Control Model.....	269
Introduction.....	269
Development of the Switching Models.....	269
Reverse-Active Region	271
Saturation	272
Junction Capacitances.....	274
Relationship Between Charge Control and Hybrid-Pi Parameters	274
Finding Junction Capacitances from the Datasheet.....	275
Manufacturers' Testing	277
Charge Control Model Examples.....	277
Emitter Switching	292
2N2222 Datasheet Excerpts.....	294
Chapter 10 Problems.....	298
References.....	303

Chapter 11: Review of Feedback Systems.....	305
Introduction and Some Early History of Feedback Control	305
Invention of the Negative Feedback Amplifier	306
Control System Basics	308
Loop Transmission and Disturbance Rejection	309
Stability	310
Routh Stability Criterion.....	311
The Phase Margin and Gain Margin Tests	314
Relationship Between Damping Ratio and Phase Margin.....	315
Loop Compensation Techniques—Lead and Lag Networks	316
Parenthetical Comment on Some Interesting Feedback Loops	317
Appendix: MATLAB Scripts	338
Chapter 11 Problems.....	342
References.....	345
Chapter 12: Basic Operational Amplifier Topologies and a Case Study	347
Basic Device Operation	347
Brief Review of LM741 Op-Amp Schematic	358
Some Real-World Limitations of Operational Amplifiers	359
Chapter 12 Problems.....	366
References.....	367
Chapter 13: Review of Current Feedback Operational Amplifiers	369
Conventional Voltage-Feedback Op-Amp and the Constant “Gain Bandwidth Product” Paradigm.....	369
Slew Rate Limitations in Conventional Op-Amps.....	371
Basic Current Feedback Op-Amp	372
Absence of Slew Rate Limit in Current Feedback Op-Amps.....	375
Manufacturer’s Datasheet Information for a Current Feedback Amplifier	379
A More Detailed Model and Some Comments on Current-Feedback Op-Amp Limitations.....	381
Chapter 13 Problems.....	382
References.....	383
Chapter 14: Analog Low-Pass Filters.....	385
Introduction.....	385
Review of Low-Pass Filter Basics	386
Butterworth Filter.....	387
Chebyshev Filter	390

Bessel Filter	395
Comparison of Responses of Different Filter Types	398
Filter Implementation.....	400
Chapter 14 Problems	414
References	416
Chapter 15: Review of Passive Components and a Case Study in PC Board Layout.....	417
Resistors	417
Comments on Surface-Mount Resistors	419
Comments on Resistor Types.....	420
Capacitors	421
Inductors	424
Discussion of Printed-Circuit Board Layout Issues.....	425
Approximate Inductance of a PC Board Trace Above a Ground Plane	428
Chapter 15 Problems	436
References.....	437
Chapter 16: Other Useful Design Techniques and Loose Ends	439
Thermal Circuits	439
Steady-State Model of Conductive Heat Transfer	440
Thermal Energy Storage	441
Using Thermal Circuit Analogies to Determine Static Semiconductor Junction Temperature	444
Mechanical Circuit Analogies.....	444
The Translinear Principle.....	449
Input Impedance of Infinitely Long Resistive Ladder	451
Transmission Lines 101	451
Node Equations and Cramer's Rule.....	455
Finding Oscillation Modes.....	458
Some Comments on Scaling Laws in Nature	463
Chapter 16 Problems	469
References.....	472
Index	475
What's on the CD-ROM?.....	480

Preface

What? Another textbook on analog circuit design? Well, this is not a textbook *per se*; rather, it is designed to be more of a design handbook for practicing engineers and students interested in learning more real-world techniques for designing and analyzing analog circuits using transistors, diodes, and operational amplifiers. The hope is that the reader will find a good mixture of theoretical techniques and also real-world design examples and test results.

This author is a practicing electrical engineer in the analog and power electronics realm who also has had the opportunity to teach at Worcester Polytechnic Institute.

The careful reader, upon review of the chapter references, will note that I have a fondness for using older references. This is due in part to the fact that the authors of these older texts and papers did not have computers available to them for circuit simulations and mathematical number-crunching. These references, in many cases, give very useful approximations, intuitive insights, and different ways of looking at difficult circuit analysis problems.

Intended Audience

This text is loosely based on a set of course notes designed for my graduate-level analog circuit design seminar offered at Worcester Polytechnic Institute. Students who take my course have already taken undergraduate-level courses covering transistors, signal processing, Bode plots and the like. Furthermore, it is the author's hope that the techniques shown in the book will be useful for practicing analog (and perhaps even *digital*!) design engineers.

Text Outline

Chapters 1 and 2 offer introductory material. Chapter 1 serves as an introduction and motivation to analog circuit design in general, with selected history thrown into the mix. Chapter 2 covers important signal-processing concepts that are used in later chapters, so that the reader will be on the "same page" as the author.

Chapters 3 to 8 cover the bipolar device physics, the bipolar junction transistor (BJT), transistor amplifiers, and approximation techniques for bandwidth estimation and switching speed analysis.

Chapter 9 covers the basics of CMOS and CMOS amplifiers. The bandwidth estimation techniques developed in earlier chapters for amplifier design work well for CMOS devices as well.

Preface

Chapter 10 covers transistor switching. How do you get a transistor to turn ON and OFF quickly, and how do you estimate that speed?

Chapter 11 is a review of feedback systems and of the Bode plot/phase margin method of designing stable feedback systems.

Chapters 12–13 cover the design, use and limitations of real-world operational amplifiers, including voltage-feedback and current-feedback op-amps.

Chapter 14 covers the basics of analog low-pass filter design, including ladder and active implementations of Butterworth, Chebyshev, elliptic and Bessel filters.

Chapter 15 covers real-world design issues such as PC board layout rules-of-thumb and the use and limitations of passive components.

Chapter 16 is a potpourri of useful design techniques and tricks that don't fit into the other chapters.

Throughout the text, some illustrative analysis problems and MATLAB® and PSPICE design examples are haphazardly sprinkled.

Acknowledgments

The author would like to acknowledge the learned professors and teaching assistants at the Massachusetts Institute of Technology who taught him many of the techniques shown in this book. They include Prof. Jim Roberge, Prof. Bill Siebert, Prof. Bill Peake, Prof. Marty Schlecht, Leo Casey, Tom Lee, Prof. Harry Lee, Prof. Campbell Searle, Prof. Amar Bose, and Prof. Dick Thornton.

This text has evolved from courses the author taught over the past years at Worcester Polytechnic Institute. Therefore, further thanks go to W.P.I. and its students and faculty who directly and indirectly contributed to this text. The author also acknowledges the indirect contributions of his W.P.I. students who, through their probing questions and careful reading of the course notes, have identified numerous typographical errors and half-truths, which hopefully have been fully expunged from this edition.

The author gratefully acknowledges the patience and technical support provided by the staff at Elsevier. The author also wishes to thank my friend and colleague Dr. Alexander Kusko, who reviewed the manuscript and offered many useful suggestions. Also thanks to my friend Dr. Jeff Roblee, who read the manuscript from a mechanical engineering point-of-view and offered many useful suggestions.

Plots were created using MATLAB and the Microsim Student Version of PSPICE, version 8.0.

Marc T. Thompson
Harvard Labs
Harvard, Massachusetts
October, 2005

Introduction and Motivation

In This Chapter

- *This chapter serves as an introduction to the philosophy and topical coverage of this book. A very brief history of transistor development, invention of the analog integrated circuit (IC) and operational amplifier advances are given.*

The Need for Analog Designers

There is an inexorable trend in recent years to “go digital”—in other words, to do more and more signal processing in the digital domain due to a purported design flexibility. However, the world is an analog place and the use of analog processing allows electronic circuits to interact with the physical world. Not discounting the importance of digital signal processing (DSP) and other digital techniques, there are many analog building blocks such as operational amplifiers, transistor amplifiers, comparators, A/D and D/A converters, phase-locked loops and voltage references (to name just a few) that are still used and will be used far into the future. Therefore, there is a continuing need for course development and education covering basic and advanced principles of analog circuit design.

One reason why analog electronic circuit design is so interesting is the fact that it encompasses so many different disciplines. Here’s a partial “shopping list,” in no particular order, of disciplines encompassed by the broad field of analog circuit design:

- *Analog filters:* Discrete or ladder filters, active filters, switched capacitor filters, crystal filters.
- *Audio amplifiers:* Power op-amps, output (speaker driver) stages.
- *Oscillators:* Including LC, crystal, relaxation and feedback oscillators, phase-locked loops, video demodulation.
- *Device fabrication and device physics:* Metal oxide semiconductor field effect transistors (MOSFETs), bipolar transistors, diodes, insulated gate bipolar transistors (IGBTs), silicon-controlled rectifiers (SCRs), MOS-controlled thyristors (MCTs), etc.
- *IC fabrication:* Operational amplifiers, comparators, voltage references, PLLs, etc.
- *Analog-to-digital interface:* A/D and D/A, voltage references.

- *Radio frequency circuits*: RF amplifiers, filters, mixers and transmission lines; cable TV.
- *Controls*: Control system design and compensation, servomechanisms, speed controls.
- *Power electronics*: This field requires knowledge of MOSFET drivers, control system design, PC board layout, and thermal and magnetic issues; motor drivers; device fabrication of transistors, MOSFETs (metal oxide semiconductor field effect transistors), IGBTs (insulated gate bipolar transistors), and SCRs (silicon-controlled rectifiers).
- *Medical electronics*: Instrumentation (EKG, NMR), defibrillators, implanted medical devices.
- *Simulation*: SPICE and other circuit simulators.
- *PC board layout*: This requires knowledge of inductance and capacitive effects, grounding, shielding and PC board design rules.

Since we do live in a world where more and more digital processing is taking place, analog designers must also become comfortable with digital-processing concepts so that we can all work together. In the digital world, some subsystem designs are based on analog counterparts. When designing a digital filter, one often first designs an analog prototype and then through an analog-to-digital transformation the filter is converted to the digital domain. For example, a bilinear transformation may be used where a filter designed in the s -domain (analog, using inductors, capacitors, and/or active elements) is transformed to a filter in the z -domain (digital, with gain elements and delays).

This technique stems in part from the fact that designers are in general more comfortable working in the analog domain when it comes to filtering. It's very easy to design a second-order analog Butterworth filter (you can find the design in any number of textbooks or analog filter cookbooks) but the implementation in the digital domain requires additional steps or other simulation tools.

Also, at sufficiently high frequencies, a digital transmission line or a high-speed signal trace on a PC board must be treated as a distributed analog system with traveling waves of voltage and current. Increasing density of digital integrated circuits and faster switching speeds are adding to the challenges of good PC board design due to extra power requirements and other issues such as ground bounce.

The bottom line is... it behooves even digital designers to know something about analog design.

Some Early History of Technological Advances in Analog Integrated Circuits

The era of semiconductor devices can arguably be traced back as far as Dr. Julius Lilienfeld, who has several U.S. patents giving various MOS structures (**Figure 1-1**). In three patents, Dr. Lilienfeld gave structures of the MOSFET, MESFET and other MOS devices.

We entered the bipolar transistor semiconductor era over 50 years ago with early work in solid-state physics and the invention of the bipolar transistor, and significant technological advances in analog circuit design and device fabrication are still being made. In 1947–1948, Bardeen, Brattain and Shockley demonstrated the first bipolar transistors (**Figure 1-2**).¹

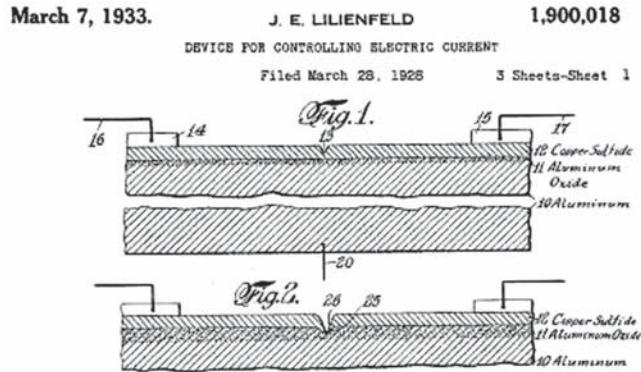


Figure 1-1: Excerpt from Lilienfeld's U.S. patent 1,900,018² (1933).

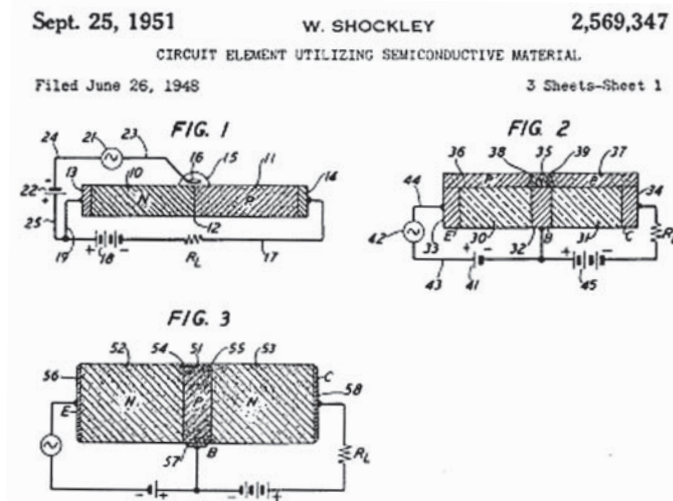


Figure 1-2: Excerpt from Shockley's U.S. patent 2,569,347 (1951).

¹ See U.S. Patent #2,569,347, "Circuit Element Utilizing Semiconductive Materials," issued September 25, 1951 to William Shockley. Bardeen, Brattain and Shockley shared the 1956 Nobel prize in physics for their discoveries related to the transistor. An excellent description of semiconductor transistor physics is given in Shockley's Nobel lecture "Transistor Technology Evokes New Physics," dated December 11, 1956.

² Lilienfeld had three patents in succession covering basic MOS transistor structures.

Chapter 1

The first integrated circuits were produced around 1959 by teams at Fairchild Semiconductor and Texas Instruments (**Figure 1-3**). TI claims invention of the integrated circuit, with J. S. Kilby's U.S. patent "Miniaturized Electronic Circuits" #3,138,743, filed Feb. 6, 1959. Workers at Fairchild filed for a patent on the first planar IC (arguably more easily manufactured than the TI invention) shortly after; see R. N. Noyce, "Semiconductor Device-and-Lead Structure," U.S. patent # 2,981, 877, filed July 30, 1959.

These ICs had minimum feature sizes of around 125 micrometers. Since then, device geometries have gotten smaller and smaller with the invention and rapid improvements in the integrated circuit (IC). Moore's Law, named for Fairchild and Intel founder Gordon Moore, predicts that the density of transistor packaging in integrated circuits doubles approximately every 18 months, a trend that has proven to be remarkably accurate over the past 30 years.

At the time of this writing,³ IC manufacturers are using 90-nanometer CMOS manufacturing processes, and smaller transistor sizes are anticipated. The smaller size allows the packaging of more and more complicated structures in a given die area. Researchers⁴ are also actively working on three-dimensional integrated circuit structures in an attempt to pack more and more functionality into a given die volume.

After the invention of the integrated circuit around 1958–1959 by workers at Texas Instruments and Fairchild, the first integrated-circuit operational amplifiers were introduced in the

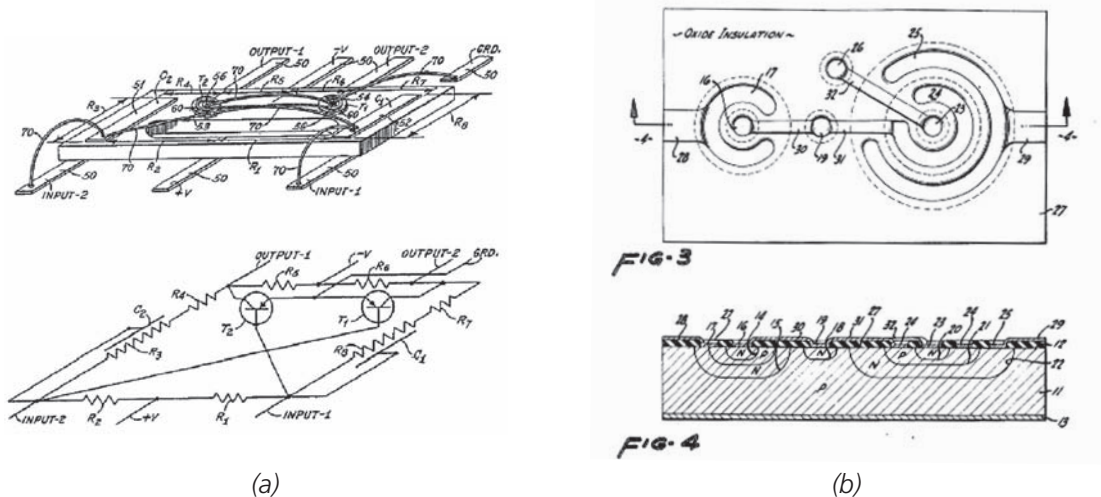


Figure 1-3: Diagrams from competing IC patents⁵ from Texas Instruments (a) and Fairchild (b).

³ Fall, 2003; note that in 1983 a typical minimum linewidth in ICs was 1.5 micrometers (1500 nanometers). Manufacturers are aiming for 65 and even 50 nanometer gate lengths. See Gordon Moore's paper "The Role of Fairchild in Silicon Technology in the Early Days of 'Silicon Valley' " where the history of Fairchild IC development is recounted.

⁴ See, e.g., Matrix Semiconductor, Inc. and "A Vertical Leap for Microchips" by Thomas H. Lee.

⁵ Full text and images of patents are available from the U.S. Patent office, <http://www.uspto.gov>.

early to mid-1960s. The first commercially successful op-amps were the Fairchild μ A709 (1965) and the National LM101 (1967), designed by the legendary analog wizard Bob Widlar.⁶ These devices had a voltage offset of a few millivolts and a unity-gain bandwidth of around a megahertz and required external components for frequency compensation. Soon after (1968), the ubiquitous Fairchild μ A741, the industry's first internally compensated op-amp, was introduced and became a bestseller. In the 741, a 30-picofarad compensating capacitor was integrated onto the chip using metal-oxide technology. It was somewhat easier to use than the LM101 because this compensating capacitor was added internal to the IC.⁷ The corresponding price reductions and specification improvements of the monolithic IC op-amps as compared to the earlier discrete designs (put forth, for instance by Philbrick)⁸ made these IC op-amps instant successes.

Since that time, op-amps have been designed and introduced with significantly better voltage offset and bandwidth specifications, as well as improvements in other specifications such as input current, common mode range, and the like. FET input op-amps became available in the 1970s with lower input current than their bipolar counterparts. Novel topologies such as the current-feedback op-amp have been introduced with success, for high-speed applications.⁹ Typical high-speed op-amps today have bandwidths of hundreds of megahertz.¹⁰ Power

⁶ The earlier μ A702 op-amp was designed by Widlar and introduced in 1963 by Fairchild but never achieved much commercial success. Widlar went back to the drawing board and came up with the 709 around 1965; it was the first op-amp to cost less than \$10. After a salary dispute with Fairchild, Widlar moved to National Semiconductor where he designed the LM101 and later improved the design resulting in the LM101A (1968). Details and history of the LM101 and 709 are given in the Widlar paper "Design Techniques for Monolithic Operational Amplifiers" with citation given at the end of this chapter.

⁷ The 741 does not need an external compensation capacitor as did previous op-amps such as the LM101 and the 709. The "plug and play" ease of use of the 741 apparently offsets the fact that under most applications with closed-loop gains greater than 1 the device is over-compensated. More details on op-amp topologies are given in a later chapter in this book. Details and history of the 709, LM101 and 741 op-amps are also given in Walt Jung's *IC Op-Amp Cookbook*, 3rd edition, pp. 75–98.

⁸ For instance, the Philbrick K2-W op-amp, made with discrete components (vacuum tubes!), and sold from 1951 to 1971. It had a small signal bandwidth of around 300 kHz and an open-loop gain of 10,000 or so. The units were priced at around \$22. See the article by Bob Pease, "What's all this K2-W Stuff, Anyway?" Philbrick also made the P2, a low input current discrete operational amplifier built with a handful of transistors and other discrete components, and priced at around \$200. See, "The Story of the P2—The First Successful Solid-State Operational Amplifier with Picoampere Input Currents" by Bob Pease, found in *Analog Circuit Design Art Science and Personalities*, edited by Jim Williams.

⁹ The current-feedback op-amp does not have constant gain-bandwidth product as does the standard voltage feedback op-amp.

¹⁰ See, e.g., National's LM6165 with a gain-bandwidth product (GBP) of 725 MHz, the Linear Technology LT1818 with GBP = 400 MHz, or the Analog Devices AD8001 with GBP = 600 MHz.

op-amps¹¹ exist that can drive speakers or other heavy resistive or inductive loads with several amperes of load current. Low-power op-amps with sub-milliwatt standby power dissipation are now commonplace. Rail-to-rail op-amps are now available.

These advances have opened new applications and product markets for devices based on analog and digital signal processing. Currently, cellular telephone, cable television, and wireless internet technologies are driving the business in RF analog circuit design and miniature handheld power electronics. Low-power devices enable the design of battery-powered devices with long battery life.

Digital vs. Analog Implementation: Designer's Choice

In many instances, functions that might be implemented in the digital domain would be difficult, costly and power-hungry to implement as compared to a relatively simple analog counterpart. For instance, consider the design of a logarithmic amplifier. One can exploit the well-known logarithmic/exponential voltage-current relationship¹² of a bipolar transistor operated in the forward-active region, as given by:

$$\begin{aligned} I_C &\approx I_S e^{\frac{qV_{BE}}{kT}} \\ V_{BE} &\approx \frac{kT}{q} \ln \left(\frac{I_C}{I_S} \right) \end{aligned} \quad [1-1]$$

This relationship holds over many orders of magnitude of transistor collector current. Therefore, one can use a transistor PN junction to implement a low-cost logarithmic amplifier (**Figure 1-4**). The input-output transfer function of this circuit, assuming an ideal transistor and op-amp, is:

$$v_o = -\frac{kT}{q} \ln \left(\frac{v_i}{RI_S} \right) \quad [1-2]$$

This circuit provides an output voltage that is proportional to the natural logarithm of the input voltage. An implementation in the digital domain would be considerably more involved. The same principles can be applied to do analog multiplication.¹³

The logarithmic relationship between voltage and current in a bipolar transistor can be used to implement analog multipliers, dividers and square root circuits. Let's look at the circuit of

¹¹ One example is the National LM12, also designed by Bob Widlar. An excellent IEEE paper discussing the design of the LM12 is "A Monolithic Power Op Amp" with citation at the end of this chapter.

¹² One can also exploit this exponential relationship to design analog multipliers, such as the "Gilbert Cell," and analog dividers and square and cube root circuits.

¹³ See, for instance, the translinear principle.

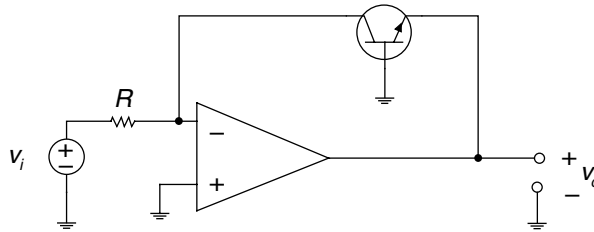


Figure 1-4: Simple logarithmic amplifier.

Figure 1-5. By using the translinear principle (discussed later in this book) we can find the relationship between the various transistor collector currents:

$$I_{C1}I_{C2} = I_{C3}I_{C4} \quad [1-3]$$

This means we can express the output current I_o as:

$$I_o = \sqrt{I_1 I_2} \quad [1-4]$$

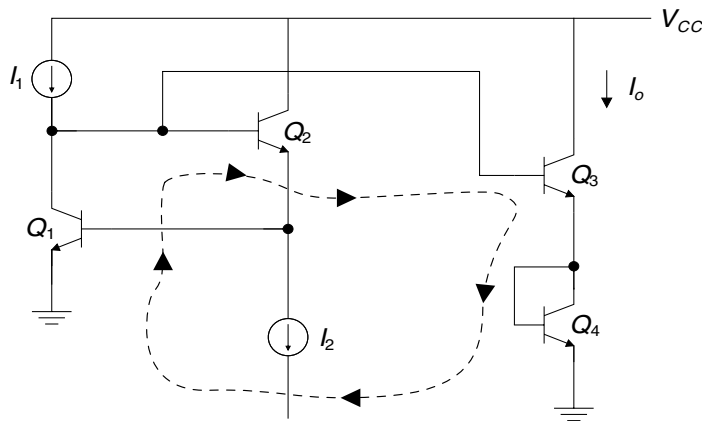
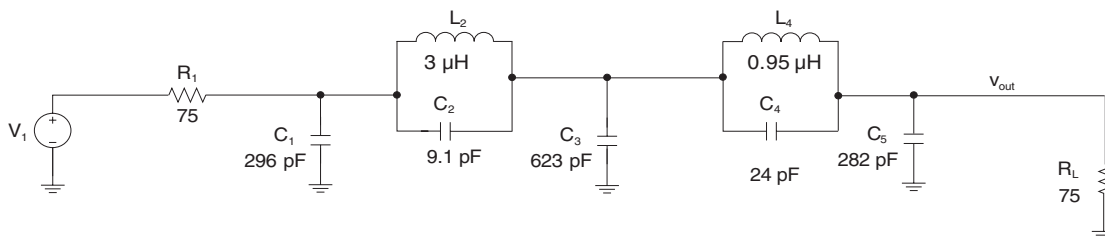


Figure 1-5: A translinear circuit, where the output current is equal to the square root of the product of the two inputs.

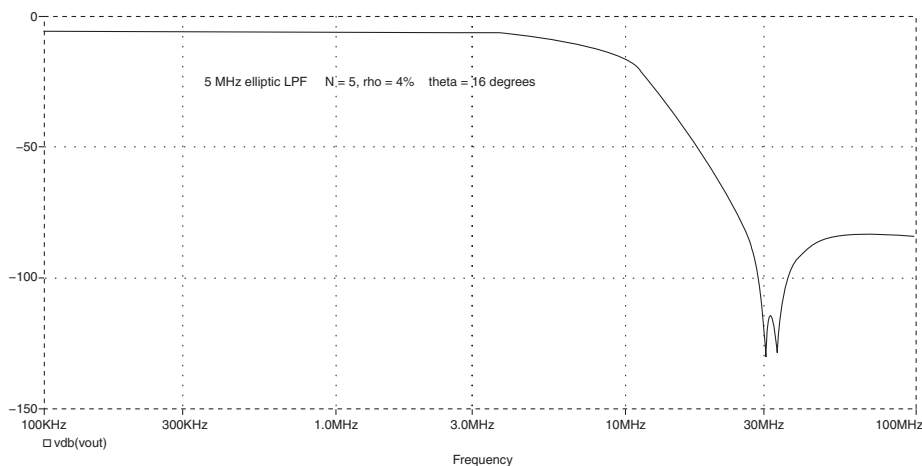
Now, consider the design of a fifth-order elliptic low-pass filter with a cutoff frequency of 5 megahertz. This is a typical specification for a video low-pass filter. This filter could be implemented in the digital domain with discrete hardware or a digital signal processor. You can implement this sharp-cutoff filter¹⁴ with just a handful of discrete components (**Figure 1-6**). Note that in this filter the source and termination resistances are each 75Ω ,

¹⁴ Elliptic filters are commonly used in analog video filtering for anti-aliasing where a very sharp cutoff transition band is required. Elliptic and other ladder filter designs are in tabulated form in Anatol Zverev's *Handbook of Filter Synthesis*. And, yes, you can build practical analog filters using inductors.

corresponding to the characteristic impedance of a typical video BNC cable. Again, an implementation in the digital domain would be significantly more complicated, especially if a high-frequency cutoff is required.



(a)



(b)

Figure 1-6: Fifth order elliptic ladder filter with 5 MHz cutoff frequency.
(a) Circuit. (b) Frequency response.

So, Why Do We Become Analog Designers?

One possible answer to this hypothetical question is to note that in any given analog design problem there is not one absolute, unique and correct answer, or “perfect” design. As a matter of fact, if you think that you have arrived at the unique, perfect solution in the analog domain you are undoubtedly mistaken. In the analog design space, there are infinities of possibilities in how to implement a given function. The challenge, and eventually (hopefully!) the reward to the analog designer, is to meet these requirements in a given design space meeting cost, size, and/or performance constraints.

Note on Nomenclature in this Text

In this text, there is a nomenclature that is used with regard to signals. In general, a transistor terminal voltage in an operating amplifier has a DC operating point and a small-signal variation about that operating point. The nomenclature used in the case of a transistor base-emitter voltage is:

$$v_{BE} = V_{BE} + v_{be} \quad [1-5]$$

where v_{BE} (small “v” and capital “BE”) is the total variable, V_{BE} (capital “V” and capital “BE”) is the DC operating point, and v_{be} (small “v” and small “be”) is the small-signal variation.

Note on Coverage in this Book

It is impossible to cover all aspects of analog design in a single textbook. Rather, in this text I have attempted to provide a potpourri of important techniques, tricks and analysis tools that I have found useful in designing real-world analog circuits. Where necessary, mathematical derivations for theoretical techniques are given. In other areas, intuitive techniques and analogies are used to “map” solutions from one design domain to the analog design domain, hopefully bypassing extensive mathematical derivations.

There are several important subjects that are conspicuously not covered in this text, including:

- Noise
- JFET amplifiers
- Switched capacitor filters

Other references, including textbooks and scholarly journals, have been provided at the end of each chapter for the reader to explore the topics in more depth. The author has provided explanatory notes including opinions with some of the citations.

It is assumed that the reader has a familiarity with Laplace transforms, pole-zero plots, Bode plots, the concept of system step response, and a basic understanding of differential equations. Chapter 2 of this book reviews these signal processing basics. These fundamentals are needed to foster understanding of the more advanced topics in later chapters.

References

- Analog Devices, *Nonlinear Circuits Handbook*, Analog Devices, 1976. *Good coverage of logarithmic amplifiers and other nonlinear analog circuits.*
- Bardeen, J., and Brattain, W. H., “The Transistor, A Semiconductor Triode,” *Physical Review*, vol. 74, no. 2, pp. 230–231, July 15, 1948, reprinted in *Proceedings of the IEEE*, vol. 86, no. 1, January 1998, pp. 29–30.
- Bondyopadhyay, Probir, “In the Beginning,” *Proceedings of the IEEE*, vol. 86, no. 1, January 1998, pp. 63–77.
- , “W = Shockley, the Transistor Pioneer—Portrait of an Inventive Genius,” *Proceedings of the IEEE*, vol. 86, no. 1, January 1998, pp. 191–217.
- Brinkman, William, Haggan, Douglas, and Troutman, William, “A History of the Invention of the Transistor and Where It Will Lead Us,” *IEEE Journal of Solid-State Circuits*, vol. 32, no. 12, December 1997, pp. 1858–1865.
- Brinkman, William, “The Transistor: 50 Glorious Years and Where We Are Going,” *1997 IEEE International Solid-State Circuits Conference*, Feb. 6–8, 1997, pp. 22–26.
- Fullagar, Dave, “A New High Performance Monolithic Operational Amplifier,” *Fairchild Application Brief*, May 1968. *Description of the 741 op-amp by its designer.*
- Jung, Walter, *IC Op-Amp Cookbook*, 3rd edition, SAMS, 1995.
- Kilby, Jack, “The Integrated Circuit’s Early History,” *Proceedings of the IEEE*, vol. 88, no. 1, January 2000, pp. 109–111.
- Lee, Thomas H., “A Vertical Leap for Microchips,” *Scientific American*, January 13, 2002. *This note describes Matrix Technologies’ efforts and motivation for producing 3D integrated circuits.*
- Manglesdorf, C., “The Changing Face of Analog IC Design,” *IEEE Transactions on Fundamentals*, E85-A, no. 2, February 2002, pp. 282–285.
- , “The Future Role of the Analog Designer,” *ISSCC 93*, session WE3, pp. 78–79.
- Melliar-Smith, C. Mark, Borrus, Michael G., Haggan, Douglas, Lowrey, Tyler, Sangiovanni-Vincentelli, Alberto, and Troutman, William, “The Transistor: An Invention Becomes a Big Business,” *Proceedings of the IEEE*, vol. 86, no. 1, January 1998, pp. 86–110.
- Moore, Gordon, “The Role of Fairchild in Silicon Technology in the Early Days of “Silicon Valley,” *Proceedings of the IEEE*, vol. 86, no. 1, January 1998, pp. 53–62.
- , “Cramming More Components onto Integrated Circuits,” *Electronics*, April 19, 1965, pp. 114–117. *The original paper where “Moore’s Law” was first published.*
- National Semiconductor, “Log Converters,” *National Application Note AN-30*.
- , *Linear Applications Handbook*.
- Pearson, G. L., and Brattain, W. H., “History of Semiconductor Research,” *Proceedings of the IRE*, vol. 43, no. 12, December 1955, pp. 1794–1806.
- Pease, Bob, “What’s All This K2-W Stuff, Anyway?,” *Electronic Design*, January 6, 2003.

- , “The Story of the P2—The First Successful Solid-State Operational Amplifier with Picoampere Input Currents,” found in *Analog Circuit Design: Art, Science and Personalities*, pp. 67–78, edited by Jim Williams, Butterworth-Heinemann, 1991. *A history and description of the K2-W operational amplifier; a discrete op-amp sold in the 50s and 60s by Philbrick, and built with vacuum tubes, and the P2, a discrete-transistor Philbrick operational amplifier.*
- Pease, Robert, *Troubleshooting Analog Circuits*, Butterworth-Heinemann, 1991.
- Perry, T. S., “For the Record: Kilby and the IC,” *IEEE Spectrum*, vol. 25, no. 13, December 1988, pp. 40–41. *Description of the history of Kilby’s invention.*
- Sah, Chih-Tang, “Evolution of the MOS Transistor—From Conception to VLSI,” *Proceedings of the IEEE*, vol. 76, no. 10, October 1988, pp. 1280–1326
- Schaller, Robert, “Moore’s Law: Past, Present and Future,” *IEEE Spectrum*, vol. 34, no. 6, June 1997, pp. 52–59.
- Shockley, William, “Transistor Technology Evokes New Physics,” *1956 Physics Nobel Prize Lecture*, December 11, 1956.
- , “Transistor Electronics—Imperfections, Unipolar and Analog Transistors,” *Proceedings of the IEEE*, vol. 85, no. 12, December 1997, pp. 2055–2080.
- , “Theory of p-n Junctions in Semiconductors and p-n Junction Transistors,” *Bell System Technical Journal*, vol. 28, no. 7, July 1949, pp. 436–489.
- , “Electrons, Holes and Traps,” *Proceedings of the IRE*, vol. 46, no. 6, June 1958, pp. 973–990.
- Siebert, William McC., *Circuits, Signals and Systems*, McGraw-Hill, 1986.
- Small, James, “General-Purpose Electronic Analog Computing: 1945–1965,” *IEEE Annals of the History of Computing*, vol. 15, no. 2, 1993, pp. 8–18.
- Soloman, James, “A Tribute to Bob Widlar,” *IEEE Journal of Solid State Circuits*, vol. 26, no. 8, August 1991, pp. 1087–1089. *Tribute and anecdotes about Bob Widlar.*
- Sporck, Charles, and Molay, Richard L., *Spinoff: A Personal History of the Industry that Changed the World*, Saranac Lake Publishing, 2001.
- Sugii, Toshihiro, Watanabe, Kiyoshi, and Sugatani, Shinji, “Transistor Design for 90-nm Generation and Beyond,” *Fujitsu Science and Technology Journal*, vol. 39, no. 1, June 2003, pp. 9–22.
- United States Patent Office, website: <http://www.uspto.gov>.
- Warner, Raymond, “Microelectronics: Its Unusual Origin and Personality,” *IEEE Transactions on Electron Devices*, vol. 48, no. 11, November 2001, pp. 2457–2467.
- Widlar, Robert J., “Design Techniques for Monolithic Operational Amplifiers,” *IEEE Journal of Solid State Circuits*, vol. SC-4, no. 4, August 1969, pp. 184–191. *Describes in detail the designs of the 709 and LM101A op-amps.*
- , “Monolithic Op Amp—The Universal Linear Component,” *National Semiconductor Linear Applications Handbook*, Application Note no. AN-4.

Widlar, Robert J., and Yamatake, Mineo, “A Monolithic Power Op-Amp,” *IEEE Journal of Solid State Circuits*, vol. 23, no. 2, April 1988. A detailed description of the design and operation of the LM12, an integrated circuit power op-amp designed in the 1980s. Applications of the LM12 are given in National application note AN-446, “A 150W IC Op Amp Simplifies Design of Power Circuits,” also written by Widlar and Yamatake.

Williams, Arthur B., and Taylor, Fred J., *Electronic Filter Design Handbook LC, Active and Digital Filters*, 2nd edition, McGraw-Hill, 1988.

Williams, Jim, Editor, *The Art and Science of Analog Circuit Design*, Butterworth-Heinemann, 1998.

———, *Analog Circuit Design: Art, Science and Personalities*, Butterworth-Heinemann, 1991. These books edited by Jim Williams present a potpourri of design tips, tricks, and personal histories by some of the most noteworthy analog designers in the world.

Zverev, A., *Handbook of Filter Synthesis*, John Wiley, 1967. An excellent overview of analog ladder filter design with special emphasis on Butterworth, Bessel, Chebyshev and elliptic implementations. Lowpass to highpass and lowpass to bandpass transformations are also discussed in detail.

U.S. Patents¹⁵

Bardeen, J., and Brattain, W. H., “Three-Electrode Circuit Element Utilizing Semiconductive Materials,” U.S. Patent # 2,524,035, filed June 7, 1948; issued October 3, 1950.

Kilby, J. S., “Miniaturized Electronic Circuits,” U.S. Patent # 3,138,743, filed February 6, 1959; issued June 23, 1964.

Lilienfeld, J., “Method and Apparatus for Controlling Electric Currents,” U.S. Patent # 1,745,175, filed October 8, 1926; issued January 28, 1930; “Amplifier for Electric Currents,” U.S. Patent # 1,877,140, filed December 8, 1928; issued September 13, 1932; “Device for Controlling Electric Current,” U.S. Patent # 1,900,018, filed March 28, 1928; issued March 7, 1933.

Noyce, R. N., “Semiconductor Device-and-Lead Structure,” U.S. Patent # 2,981,877, filed July 30, 1959; issued April 25, 1961.

Shockley, W., “Circuit Element Utilizing Semiconductive Materials,” U.S. Patent # 2,569,347, filed June 17, 1948; issued October 3, 1950; “Semiconductor Amplifier,” U.S. Patent # 2,502,488, filed September 24, 1948; issued April 5, 1950, reprinted in *Proceedings of the IEEE*, vol. 86, no. 1, January 1998, pp. 34–36.

¹⁵ All patents listed are available on the U.S. Patent and Trademark website: www.uspto.gov.

Review of Signal-Processing Basics

In This Chapter

- In this chapter, the basics of signal processing and analysis are covered. The important tools offered in this chapter such as transfer functions in the Laplace domain, the concept of poles and zeros, step and impulse responses, and Bode plots are reviewed and needed by the reader in later chapters.

Review of Laplace Transforms, Transfer Functions and Pole-Zero Plots

The transfer function and pole-zero plot of any linear time-invariant (LTI) system can be found by replacing all electronic components with their impedance expressed in the Laplace domain. For instance, the transformation from the circuit domain to the Laplace (or s) domain is made by making the following substitutions of circuit elements:

Circuit domain	Laplace (s) domain
Resistance, R	R
Inductance L	Ls
Capacitance C	$\frac{1}{Cs}$

The resultant transformed circuit also expresses a differential equation; the differential equation can be found by making the substitution:

$$s \Rightarrow \frac{d}{dt} \quad [2-1]$$

Chapter 2

This transfer function of any lumped LTI system always works out to have polynomials in the Laplace¹ variable s . For instance, a typical transfer function with multiple poles and zeros has the form:

$$H(s) = \frac{a_n s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + 1}{b_m s^m + a_{m-1} s^{m-1} + \cdots + b_1 s + 1} \quad [2-2]$$

The values of s where the denominator of $H(s)$ becomes zero are the *poles* of $H(s)$, and the values of s where the numerator of $H(s)$ becomes zero are the *zeros* of $H(s)$. In this case, there are n zeros and m poles in the transfer function.

The values of s where poles and zeros occur can be either real or imaginary. Real-axis poles result in step responses expressed as simple exponentials (of form $e^{-t/\tau}$) without overshoot and/or ringing. Complex poles always exist in pairs and can result in overshoot and ringing in the transient response of the circuit if damping is sufficiently low.

Consider the simple circuit of **Figure 2-1a**, with two resistors and a capacitor. We can transform this circuit to the Laplace domain by recognizing that the impedance of a capacitor becomes $1/(Cs)$, resulting in the circuit of **Figure 2-1b**.

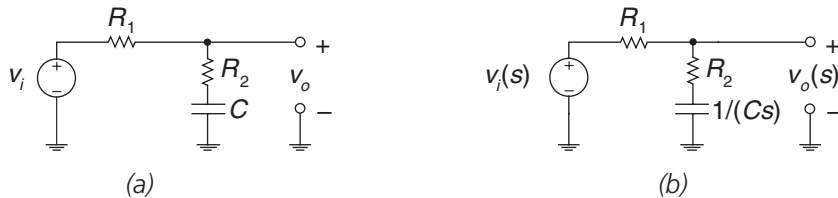


Figure 2-1: Simple RC circuit for finding system function and poles and zeros.
(a) Original circuit. (b) Circuit transformed to the Laplace domain. Note that the capacitor has been transformed into the Laplace domain resulting in an impedance of $1/Cs$.

Using **Figure 2-1b** and the resistive divider relationship, the input/output transfer function is:

$$H_1(s) = \frac{v_o(s)}{v_i(s)} = \frac{R_2 + \frac{1}{Cs}}{R_1 + R_2 + \frac{1}{Cs}} = \frac{R_2 Cs + 1}{(R_1 + R_2)Cs + 1} \quad [2-3]$$

Does this transfer function make sense? At very low frequencies approaching zero (which we can evaluate by making the substitution $s \rightarrow 0$), the transfer function is approximately:

$$H_1(s) \Big|_{s \rightarrow 0} \approx 1 \quad [2-4]$$

¹ A “lumped” circuit has any number of resistors, capacitors, inductors or dependent sources operating at frequencies low enough where wave phenomena and transmission line effects can be ignored. Remember that you can form this polynomial by taking your original circuit, and replacing each inductor with a component with impedance Ls and replacing each capacitor with a Laplace component with impedance $1/(Cs)$, and solving normally using nodal analysis or other techniques.

This makes sense, since the capacitor becomes an open-circuit at zero frequency and the entire input signal is passed through to the output. At very high frequencies, (or when $s \rightarrow \infty$) the capacitor shorts out and the transfer function is approximately:

$$H_1(s) \Big|_{s \rightarrow \infty} \approx \frac{R_2}{(R_1 + R_2)} \quad [2-5]$$

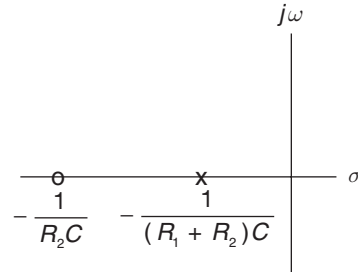
This also makes sense, since at very high frequencies this circuit looks like a simple voltage divider with R_1 and R_2 . The pole and zero of $H_1(s)$ are found by:

$$s_{pole} = -\frac{1}{(R_1 + R_2)C}$$

$$s_{zero} = -\frac{1}{R_2C} \quad [2-6]$$

The pole-zero plot for this transfer function is shown in **Figure 2-2** below.

Figure 2-2: Pole-zero plot for simple RC circuit above with a low-frequency pole and a higher-frequency zero.



This same procedure can be used for more complicated circuits with any number of resistors, capacitors, inductors and dependent sources.

First-Order System Response

Common first-order systems are the voltage-driven RC circuit (**Figure 2-3a**) and the current-driven RC circuit (**Figure 2-3b**). Topologically, these circuits result in output responses that are the same.

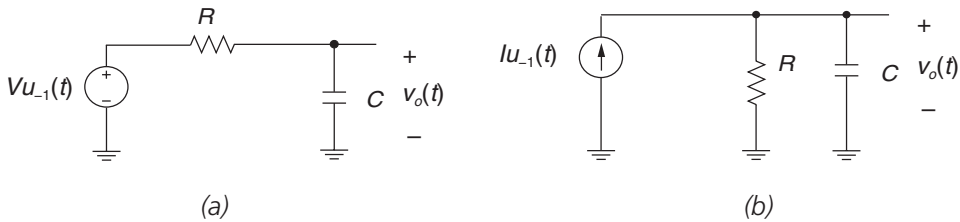


Figure 2-3: First-order RC circuits.² (a) Voltage-driven by input voltage source.
(b) Current-driven by input current source.

² Note the notation that $u_{-1}(t)$ is the unit step.

The step responses for the circuits are:

Voltage-driven RC

$$v_o(t) = V(1 - e^{-t/\tau})$$

$$i_r(t) = \frac{V}{R} e^{-t/\tau}$$

$$\tau = RC$$

[2-7]

Current-driven RC

$$v_o(t) = IR(1 - e^{-t/\tau})$$

$$i_r(t) = I e^{-t/\tau}$$

$$\tau = RC$$

The risetime of the voltage or current is defined as the time it takes the response to rise from 10% of final value to 90% of final value. For the first-order system, this risetime is given by:

$$\tau_R = 2.2\tau \quad [2-8]$$

This definition of 10–90% risetime is shown in the step response of a generic lumped first-order system in **Figure 2-4**.

The bandwidth is defined as the frequency at which the magnitude of the output response to an AC input signal drops to 70.7% (or –3dB) of the DC value. For the first-order system, the transfer function can be expressed as:

$$H(s) = \frac{1}{\tau s + 1} \quad [2-9]$$

$$\angle H(s) = -\tan^{-1}(\omega\tau)$$

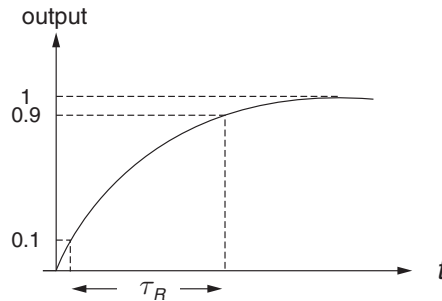


Figure 2-4: Unity step response of first-order system showing 10–90% risetime.

We find the -3dB bandwidth is:

$$\omega_h = \frac{1}{\tau} \quad [2-10]$$

$$f_h = \frac{\omega_h}{2\pi}$$

where ω_h is in radians per second and f_h is in hertz. From these relationships, we can derive the relationship between bandwidth and risetime, which is exact for a first-order system:

$$\tau_R = \frac{0.35}{f_h} \quad [2-11]$$

In **Figure 2-5**, we see the unit step response and the frequency response for a first-order system with cutoff frequency 1 radian/second.

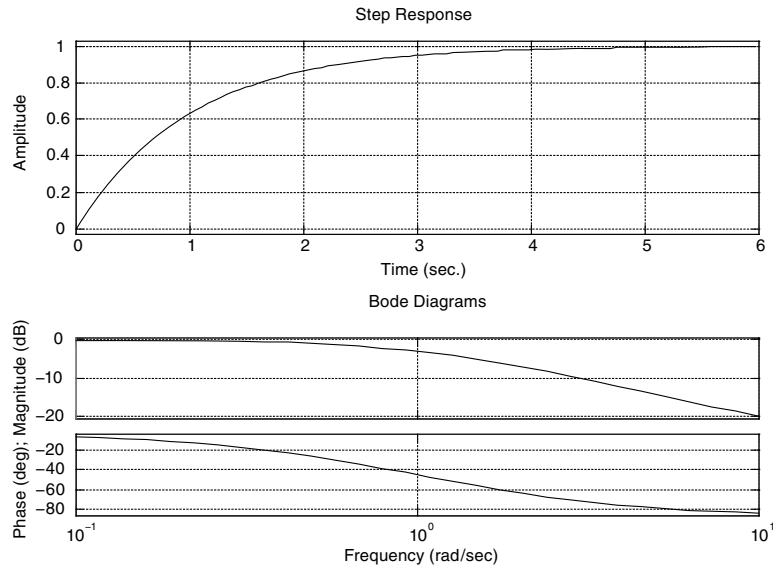


Figure 2-5: First-order RC circuits step and frequency responses for a system with cutoff frequency 1 radian/second.

First-order system approximations at low and high frequency

The transfer function of a first-order system can be expressed as:

$$H(s) = \frac{1}{\tau s + 1} \quad [2-12]$$

where τ is the “time constant” of this first-order system. The magnitude, angle and group delay³ of this transfer function are found by:

$$\begin{aligned}
 H(j\omega) &= \frac{1}{j\omega\tau + 1} \\
 |H(j\omega)| &= \frac{1}{\sqrt{(\omega\tau)^2 + 1}} \\
 \angle H(j\omega) &= -\tan^{-1}(\omega\tau) \\
 G(j\omega) &= \frac{-d\angle H(j\omega)}{d\omega} = \frac{\tau}{1 + (\omega\tau)^2}
 \end{aligned} \tag{2-13}$$

Now, let’s note how this single pole behaves at low and high frequencies (where low and high are defined in relation to the pole frequency). For high frequencies, where $\omega\tau \gg 1$, we can approximate the magnitude, phase and group delay of the transfer functions as:⁴

$$\begin{aligned}
 |H(j\omega)|_{\omega\tau \gg 1} &\approx \frac{1}{\omega\tau} \\
 \angle H(j\omega)_{\omega\tau \gg 1} &\approx -\frac{\pi}{2} + \frac{1}{\omega\tau} \\
 G(j\omega)_{\omega\tau \gg 1} &\approx \frac{1}{\omega^2\tau}
 \end{aligned} \tag{2-14}$$

For high frequencies, the magnitude of the transfer function decreases proportionally with frequency, corresponding to a -20dB/decade rolloff. The angle approaches $-\pi/2$ radians, or -90 degrees. The group delay decreases with frequency; higher-frequency components are delayed less than lower-frequency components.

For low frequencies, where $\omega\tau \ll 1$, we can approximate the transfer functions as:⁵

$$\begin{aligned}
 |H(j\omega)|_{\omega\tau \ll 1} &\approx 1 - \frac{1}{2}(\omega\tau)^2 \approx 1 \\
 \angle H(j\omega)_{\omega\tau \ll 1} &\approx -\omega\tau \\
 G(j\omega)_{\omega\tau \ll 1} &\approx \tau
 \end{aligned} \tag{2-15}$$

³ Group delay is a measure of how much time delay the frequency components in a signal undergo. Mathematically, the group delay of a system is the negative derivative of the phase with respect to omega. To find group delay for the first-order system, we make use of the identity:

$$\frac{d}{dx}(\tan^{-1} u) = \frac{1}{1+u^2} \frac{du}{dx}$$

⁴ We make use of several trigonometric identities here. For instance, the power series expansion for arctangent is $\tan^{-1}(x) = \pi/2 - 1/x + 1/(3x^3) - \dots$ for $x > 1$.

⁵ Here we make use of the power series expansion: $\frac{1}{\sqrt{1+x}} \approx 1 - \frac{x}{2}$ for $x \ll 1$.

For low frequencies, the magnitude of the transfer function is approximately unity; however, there is some finite phase shift due to the pole. For instance, at a frequency ten times lower than the pole frequency, the pole contributes approximately -0.1 radians (or -5.7 degrees) of negative phase shift. Also, at frequencies much lower than the pole frequency, the negative phase shift is approximately linear with the frequency. This means that a high-frequency pole behaves approximately as a time delay, as shown by the group delay calculation. The negative phase shift can have consequences in, for instance, feedback loops where the extra phase shift can cause oscillations.

The group delay for a first-order low-pass filter with cutoff frequency $\omega_c = 1$ radian per second is shown in **Figure 2-6**. Note that at low frequencies (low compared to the cutoff frequency) the group delay is approximately flat. Therefore, for frequencies in this region the low-pass filter behaves approximately as a constant time delay.

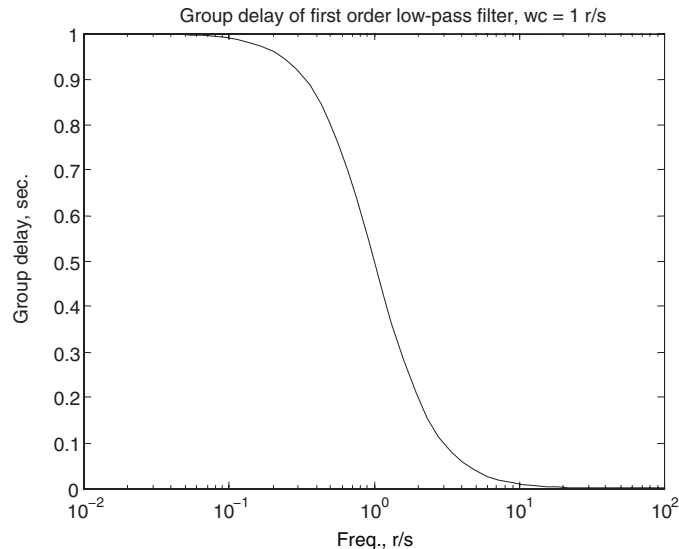


Figure 2-6: Group delay of first order low-pass filter with cutoff frequency $\omega_c = 1$ radian/s. Note that at very low frequencies ($\omega \ll \omega_c$) the group delay is approximately 1 second.

First-order system step response for short times

Let's examine a first-order step response for time short compared to the time constant of the system. We know that the step response for the first-order system of **Figure 2-7** is:

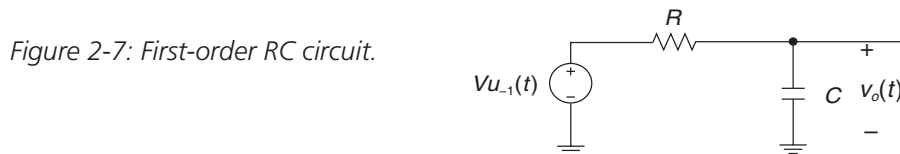


Figure 2-7: First-order RC circuit.

$$v_o(t) = V \left(1 - e^{\frac{-t}{RC}} \right) \quad [2-16]$$

where V is the amplitude of the input step. Let's assume that the input steps at time $t = 0$; let's examine the behavior of this step response for times short compared to RC . We can use the series expansion for e^x , which is:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad [2-17]$$

Therefore, for $t \ll RC$ we can approximate the step response as:

$$v_o(t) = V \left(1 - \left(1 - \frac{t}{RC} - \left(\frac{1}{2!} \right) \left(\frac{t}{RC} \right)^2 - \dots \right) \right) \approx V \left(\frac{t}{RC} \right) \text{ if } t \ll RC \quad [2-18]$$

Therefore, the step response “looks” linear in the initial time period $t \ll RC$ after $t = 0$. The capacitor current is roughly constant for $t \ll RC$:

$$i(t) = \frac{V - v_o(t)}{R} \approx \frac{V}{R} \left(1 - \frac{t}{RC} \right) \text{ if } t \ll RC \quad [2-19]$$

We can use this result to help us analyze the operation of a full-wave rectifier (**Figure 2-8**) using PSPICE.⁶ The input sinewave is a 60-Hz, 120-VRMS waveform,⁷ which has peaks at $\pm 170\text{V}$. During the positive peaks of the sinewave, D_1 and D_4 turn ON and charge up the 1000- μF load capacitor. During negative peaks of the input sinewave, D_2 and D_3 are on. Therefore, the fundamental component of the ripple of the output load voltage occurs at 120 Hz.

Note that in this circuit we have an exponential decay with RC time constant (100Ω) ($1000\ \mu\text{F}$) = 100 ms. Since this time constant is much longer than the period of the fundamental ripple frequency ($1/120\ \text{Hz} = 8.3\ \text{ms}$) the decay of the exponential “looks” approximately linear, as shown in the SPICE plot. We can very roughly calculate the amplitude of the ripple voltage on the capacitor using the following reasoning:

- The output voltage (ignoring ripple) is approximately 170V.
- The current in the load resistor is approximately $170\text{V}/100\Omega = 1.7\text{A}$
- The capacitor is discharging for (very roughly) 8.3 ms when all diodes are OFF and load current is supplied from the hold-up capacitor.
- Using $I = Cdv/dt$, we find that the ripple voltage on the capacitor is:

$$\Delta v_o \approx \frac{I_L T_{\text{discharge}}}{C} \approx \frac{(1.7)(8.3 \times 10^{-3})}{1000 \times 10^{-6}} \approx 14\text{V} \quad [2-20]$$

⁶ These plots and subsequent PSPICE analyses were done using the Microsim Evaluation PSPICE, version 8.0.

⁷ This terminology means volts root-mean-square. The RMS value of a sinewave with peaks at $\pm 170\text{V}$ is approximately 120 volts RMS. Alternate terminology would be 120 VAC, which means an AC sinewave with value 120 VRMS.

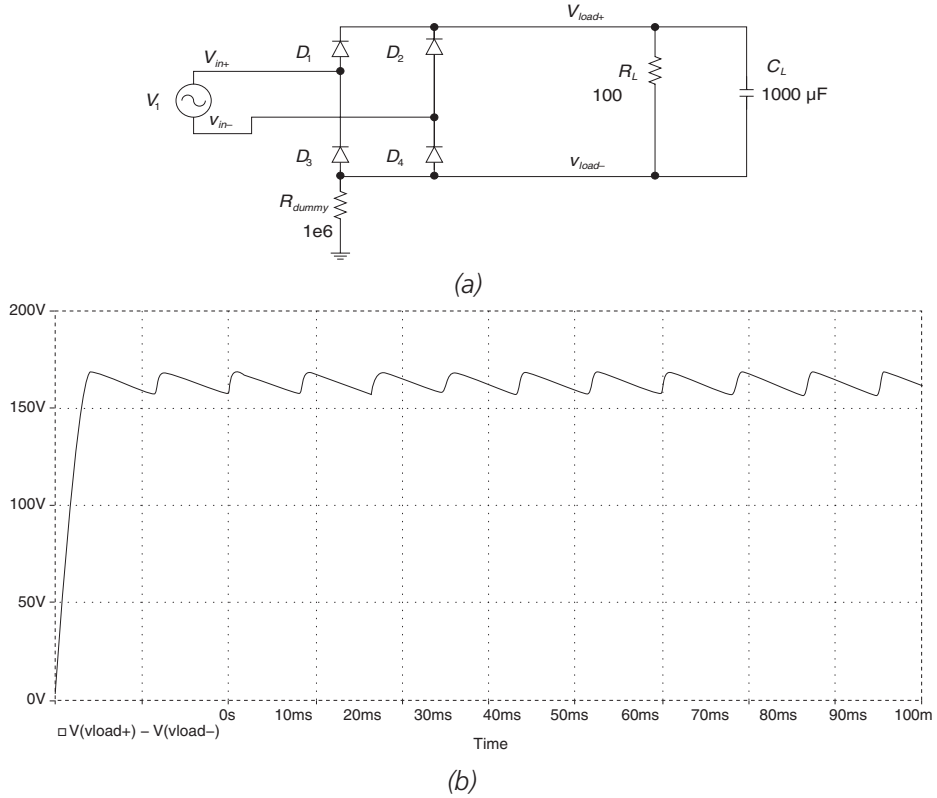


Figure 2-8: 60-Hz full-wave rectifier. (a) Circuit.⁸ (b) SPICE simulation for output voltage ($v_{load+} - v_{load-}$).

First-order system with extra high-frequency pole

How does a system behave if it has a dominant low-frequency pole and a second pole at a much higher frequency? Specifically, let's first consider a first-order system with a pole at -1 radian/second, and a transfer function:

$$H(s) = \frac{1}{(s+1)} \quad [2-21]$$

Let's next investigate how the system behaves if it has a higher-order pole added at -10 radian/second, resulting in a modified transfer function $H'(s)$:

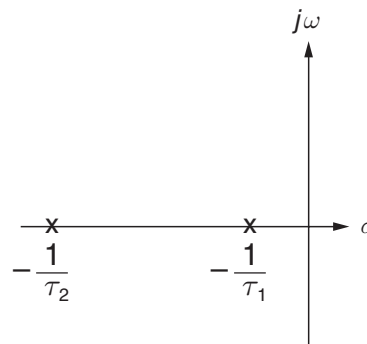
$$H'(s) = \frac{1}{(s+1)(0.1s+1)} \quad [2-22]$$

Shown following is the pole-zero plot (**Figure 2-9**) for a system $H'(s)$ with a dominant pole at -1 radian/second and a second higher-frequency pole at -10 radian/second. We expect the

⁸ Note that SPICE requires a DC path from the circuit to ground, hence the 1-M Ω dummy resistor, which doesn't significantly affect circuit operation.

step response and frequency response of this system to be dominated by the pole at -1 radian/second. For instance, we expect a 10–90% risetime of ~ 2.2 seconds, and a -3dB bandwidth of 1 radian/second.

Figure 2-9: First-order system pole-zero plot with a dominant pole at frequency $-1/\tau_1$ and a higher frequency pole at $-1/\tau_2$.



The step response (**Figure 2-10**) shows that the risetime is indeed dominated by the low-frequency pole. However, we note a subtlety: the modified system shows that in the time range 0 to ~ 3 seconds the output of the modified system lags the output of the original system. This is consistent with our previous analysis that high-frequency poles behave as time delays for frequencies compared to the pole frequency. In the case of a high-frequency pole at -10 radians/second, we expect the output of $H'(s)$ to lag that of $H(s)$ by approximately 0.1 seconds.

A comparison of the frequency response plots of a first-order system with $H'(s)$ is shown in **Figure 2-11**. We note that, at low frequencies, the frequency response is dominated by the low-frequency pole, as expected. We also note that the magnitude responses match very well, while there is some extra phase shift in the phase response due to the high-frequency pole.

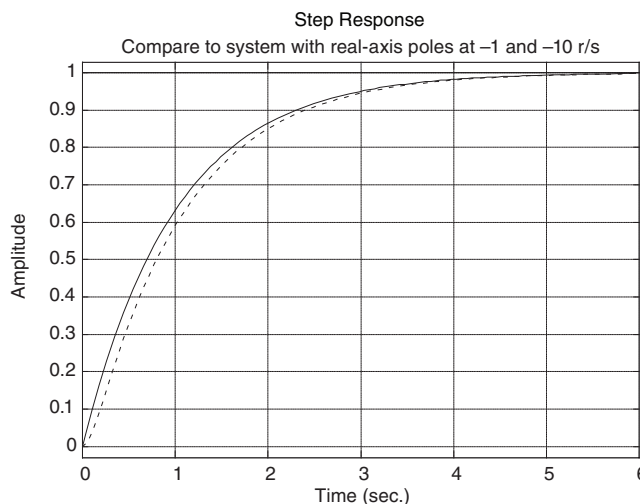


Figure 2-10: First-order system with added high-frequency pole (step response). Note that the step response of the system with the added pole lags the original system.

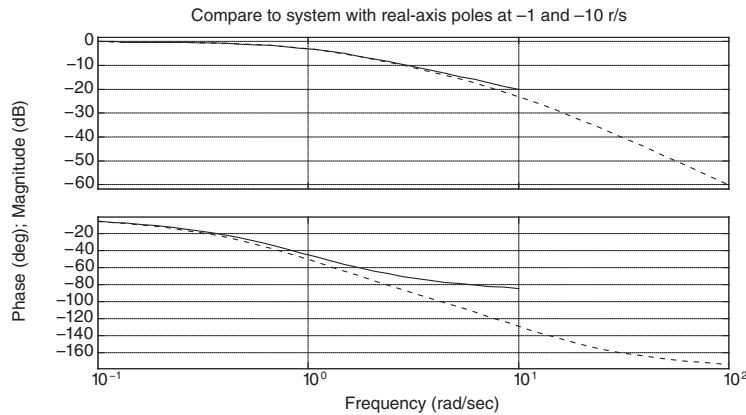


Figure 2-11: First-order system pole-zero plot with added high-frequency pole (frequency response). Note that at frequencies below that of the low-frequency pole the two magnitude transfer functions match well.

Second-Order Systems

A second-order system is one in which there are two poles. For second-order systems consisting of resistors and capacitors (without any inductors or dependent sources), the poles lie on the real axis. For this special case, there is no possibility of overshoot or ringing in the step response.

Other second-order systems can be built with coupled, lumped energy-storage mechanisms or with dependent sources, which may result in overshoot and ringing in the transient response, provided that damping is not too high. Some examples of systems that may be potentially underdamped are:

- Mass and spring
- LC circuit
- Rotational inertia and torsional spring
- RC circuits with op-amp feedback
- Magnetic suspensions

Mass-spring system

A mass-spring system is shown in **Figure 2-12**. When this system vibrates, energy is alternately stored in the kinetic energy of the mass and in the stretch of the spring. If we consider y to be the displacement of the spring from free-hanging equilibrium, the force that the spring exerts on the mass is:

$$f_y = -ky \quad [2-23]$$

Newton's law applied to the moving mass is:

$$f_y = -ky = M \frac{d^2 y}{dt^2} \quad [2-24]$$

This results in the differential equation of motion for this mass spring of:

$$M \frac{d^2 y}{dt^2} + ky = 0 \quad [2-25]$$

Let's guess a solution:⁹

$$y(t) = Y_o \sin(\omega t) \quad [2-26]$$

Putting this into the differential equation results in:

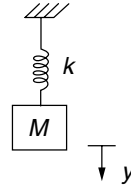
$$M(-\omega^2 Y_o \sin(\omega t)) + k(Y_o \sin(\omega t)) = 0 \quad [2-27]$$

The solution of this equation is:

$$\omega = \sqrt{\frac{k}{M}} \quad [2-28]$$

which we recognize as the oscillation frequency of a mass-spring system.

Figure 2-12: Undamped mass-spring system. The mass M (kilograms) is coupled to a spring with spring constant k (Newtons/meter). The resultant oscillation frequency is $\sqrt{k/M}$ radians/sec.



A second-order electrical system

A second-order series RLC circuit is shown in **Figure 2-13a**. Transforming to the Laplace domain results in the circuit of **Figure 2-13b**. If the resistance R is small (and we'll define what "small" means later on), this system is underdamped. The transfer function can be expressed as:

$$H(s) = \frac{v_o(s)}{v_i(s)} = \frac{\frac{1}{Cs}}{R + Ls + \frac{1}{Cs}} = \frac{1}{LCs^2 + RCs + 1} = \frac{1}{\frac{s^2}{\omega_n^2} + \frac{2\zeta s}{\omega_n} + 1} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad [2-29]$$

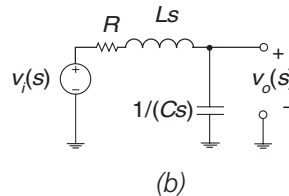
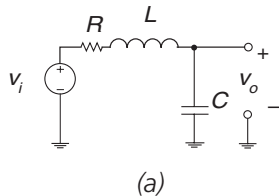


Figure 2-13: Second-order series RLC system.
(a) Circuit. (b) Circuit transformed to the Laplace domain.

⁹ We can put our intuition to work here. We know that a mass-spring system vibrates with some constant frequency. Ignoring losses in the spring and air friction, this system will vibrate forever.

This equation has been put into second-order system standard form where ω_n is the “natural frequency” of the circuit and ζ is the “damping ratio.” In the series RLC circuit, the natural frequency and damping ratio are:

$$\begin{aligned}\omega_n &= \frac{1}{\sqrt{LC}} \\ \zeta &= \frac{\omega_n RC}{2} = \frac{1}{2} \frac{R}{\sqrt{\frac{L}{C}}} = \frac{1}{2} \frac{R}{Z_o}\end{aligned}\quad [2-30]$$

The natural frequency is an indication of the relative speed of response of the system, and the damping ratio tells you how oscillatory the step response will be, and how peaky the frequency response will be. Note that the damping ratio depends on the value of the series resistor R compared to Z_o . In resonant LC circuits, this term Z_o crops up again and again and hence is called the “characteristic impedance” of the resonant circuit.

The magnitude and phase of the frequency response are given by:

$$\begin{aligned}H(j\omega) &= \frac{1}{\frac{2j\zeta\omega}{\omega_n} + \left(1 - \frac{\omega^2}{\omega_n^2}\right)} \\ |H(j\omega)| &= \frac{1}{\sqrt{\left(\frac{2\zeta\omega}{\omega_n}\right)^2 + \left(1 - \frac{\omega^2}{\omega_n^2}\right)^2}} \\ \angle H(j\omega) &= -\tan^{-1} \frac{\left(\frac{2\zeta\omega}{\omega_n}\right)}{\left(1 - \frac{\omega^2}{\omega_n^2}\right)} = -\tan^{-1} \left(\frac{2\zeta\omega\omega_n}{\omega_n^2 - \omega^2}\right)\end{aligned}\quad [2-31]$$

A plot of this transfer function magnitude and phase for varying values of damping ratio ζ is shown in **Figure 2-14**. Note that as the damping ratio decreases, the peak of the frequency response¹⁰ (which occurs at a frequency near ω_n) increases.

¹⁰ We can show that the frequency at which the transfer function has its peak is:

$$\omega_p = \omega_n \sqrt{1 - 2\zeta^2} \text{ for } \zeta < 0.707$$

This frequency is close to the natural frequency ω_n if the damping ratio is small. The magnitude of the transfer function at this peak frequency is:

$$M_p = \frac{1}{2\zeta\sqrt{1 - \zeta^2}} \text{ for } \zeta < 0.707$$

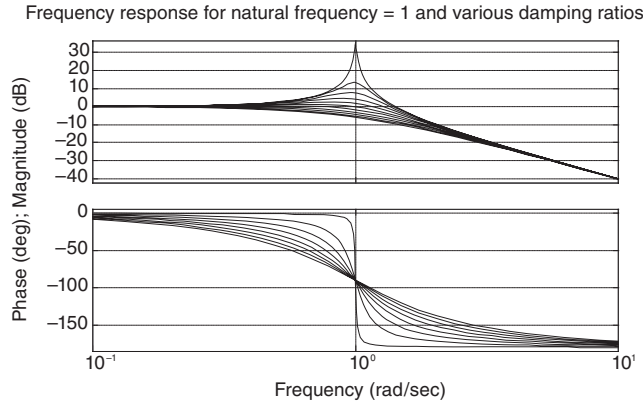


Figure 2-14: Frequency response for second-order systems, for damping ratios 0.01, 0.11, 0.21 ... 1.01; natural frequency $\omega_n = 1$. Note that for low damping there is significant peaking in the frequency response near 1 radian/second.

For R large compared to Z_o , the damping ratio is large compared to one, corresponding to the “overdamped” case. The poles of $H(s)$ in the overdamped case are found by:

$$s_{poles} = -\omega_n \left[-\zeta \pm \sqrt{\zeta^2 - 1} \right] \quad [2-32]$$

Note that as the damping ratio ζ varies, the pole locations vary as well. For zero damping (i.e., $\zeta = 0$) the poles lie on the $j\omega$ axis at $\pm j\omega_n$, corresponding to a very oscillatory response to a step excitation at this frequency. For the zero damping case, the system oscillates at the natural frequency ω_n . For critical damping (with $\zeta = 1$) the poles are both on the real axis at the same location at $s = -\omega_n$. The overdamped case occurs for $\zeta > 1$; in this case there is one pole at less than $-\omega_n$ radians per second, and a higher-frequency pole at higher than $-\omega_n$ radians per second. Pole locations for the underdamped, critically damped and overdamped cases are shown in **Figure 2-15**.

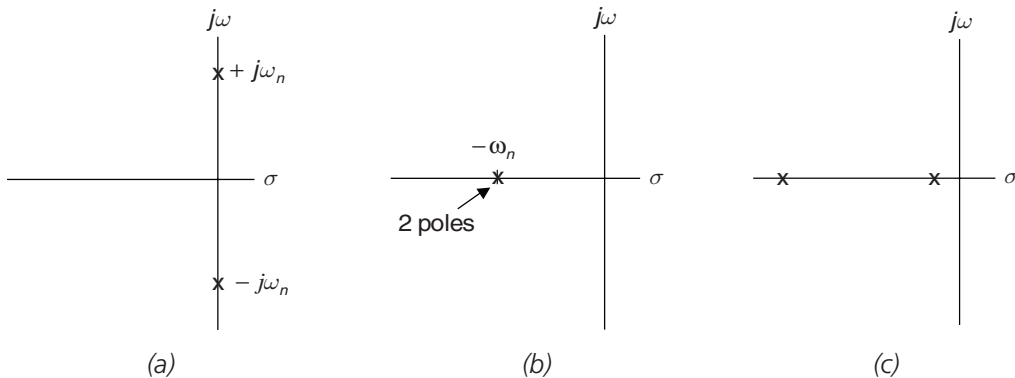


Figure 2-15: Pole-zero plot for series RLC circuit. (a) Underdamped circuit with $R \ll Z_o$. (b) Critically damped with $R = 2Z_o$. (c) Overdamped with $R \gg Z_o$.

Now, what happens if we excite this system exactly at the natural frequency, or at $\omega = \omega_n$? The response is:

$$\begin{aligned} |H(j\omega)|_{\omega=\omega_n} &= \frac{1}{2\zeta} \\ \angle H(j\omega)_{\omega=\omega_n} &= -\frac{\pi}{2} \end{aligned} \quad [2-33]$$

For low damping ratio ζ , the frequency response shows a very peaky response at the natural frequency.

Quality factor “Q”

Another commonly used metric for evaluating second-order systems is *quality factor* or “Q.” Quality factor is defined as:

$$\omega \frac{E_{\text{stored}}}{P_{\text{diss}}} \quad [2-34]$$

where E_{stored} is the peak stored energy in the system and P_{diss} is the average power dissipation. First, let’s look at a series resonant network (**Figure 2-16a**). The calculation for quality factor at resonance is:

$$\begin{aligned} \omega &= \frac{1}{\sqrt{LC}} \\ E_{\text{stored}} &= \frac{1}{2} L I_{pk}^2 \\ P_{\text{diss}} &= \frac{1}{2} I_{pk}^2 R \\ Q &= \left(\frac{1}{\sqrt{LC}} \right) \left(\frac{\frac{1}{2} L I_{pk}^2}{\frac{1}{2} I_{pk}^2 R} \right) = \frac{\sqrt{L}}{\sqrt{C}} = \frac{Z_o}{R} \end{aligned} \quad [2-35]$$

In the series resonant circuit, as resistance R increases, quality factor decreases. This is because a large resistor provides more damping.

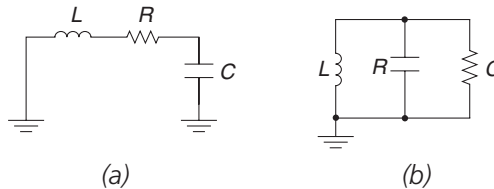


Figure 2-16: Resonant networks.
(a) Series resonant RLC network. (b) Parallel resonant.

Next, let's consider a parallel RLC network (**Figure 2-16b**). The calculation for Q at resonance is:

$$\begin{aligned}\omega &= \frac{1}{\sqrt{LC}} \\ E_{stored} &= \frac{1}{2} C V_{pk}^2 \\ P_{diss} &= \frac{1}{2} \frac{V_{pk}^2}{R} \\ Q &= \left(\frac{1}{\sqrt{LC}} \right) \left(\frac{\frac{1}{2} C V_{pk}^2}{\frac{1}{2} \frac{V_{pk}^2}{R}} \right) = \frac{R}{\sqrt{\frac{L}{C}}} = \frac{R}{Z_o}\end{aligned}\tag{2-36}$$

In the parallel resonant circuit, a small resistor provides more damping.

Now, we see that quality factor is related to damping ratio by:

$$Q = \frac{1}{2\zeta}\tag{2-37}$$

Let's rewrite the previous equation for the transfer function of a series-resonant circuit, this time using quality factor:

$$H(s) = \frac{1}{\frac{s^2}{\omega_n^2} + \frac{2\zeta s}{\omega_n} + 1} = \frac{1}{\frac{s^2}{\omega_n^2} + \frac{s}{\omega_n Q} + 1}\tag{2-38}$$

The magnitude of this transfer function is:

$$|H(s)| = \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_n^2}\right)^2 + \left(\frac{\omega}{\omega_n Q}\right)^2}}\tag{2-39}$$

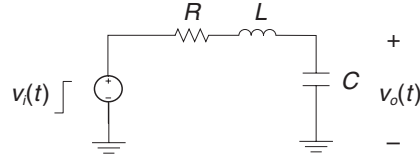
At the natural frequency ($\omega = \omega_n$), the magnitude of the transfer function is:

$$|H(s)|_{\omega=\omega_n} = Q\tag{2-40}$$

Transient response of second-order system

A second-order LCR system is shown in **Figure 2-17**. The input is a step $v_i(t)$ at time $t = 0$. Let's assume, at first, that the resistance is small enough so that the system is under-damped (i.e., that it has complex poles).

Figure 2-17: Second-order system with step input.



The step response can be expressed as:¹¹

$$v_o(t) = 1 - \frac{e^{-\zeta\omega_d t}}{\sqrt{1-\zeta^2}} \sin\left(\omega_d t + \tan^{-1}\left(\frac{\sqrt{1-\zeta^2}}{\zeta}\right)\right) \quad [2-41]$$

$$\omega_d = \omega_n \sqrt{1-\zeta^2}$$

The magnitude of the peak of the step response is:

$$P_o = 1 + e^{\frac{-\pi\zeta}{\sqrt{1-\zeta^2}}} \quad [2-42]$$

For instance, for a damping ratio of 0.2, the peak overshoot for a second-order system is approximately 1.6.

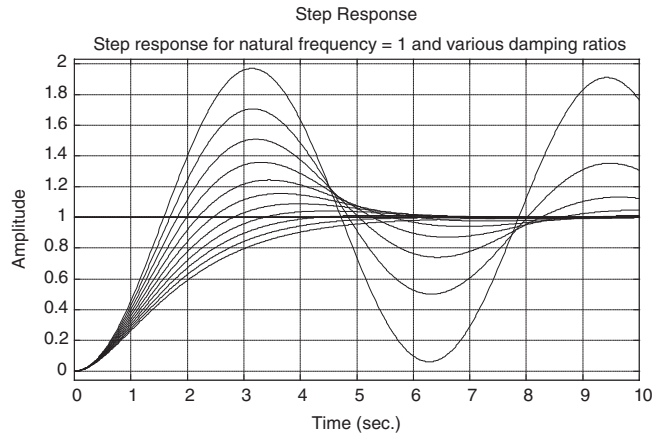
By Laplace analysis we can determine that the following parameters hold:

$$\omega_n = \frac{1}{\sqrt{LC}} \quad [2-43]$$

$$\zeta = \frac{R}{2\sqrt{\frac{L}{C}}}$$

So, for resistance small compared to $\sqrt{L/C}$ the system is underdamped. The step response for a second-order system with $\omega_n = 1$ and damping ratio varying from 0.01 to 1.01 is shown in **Figure 2-18**.

Figure 2-18: Step response for second-order systems, for damping ratios 0.01, 0.11, 0.21 ... 1.01.



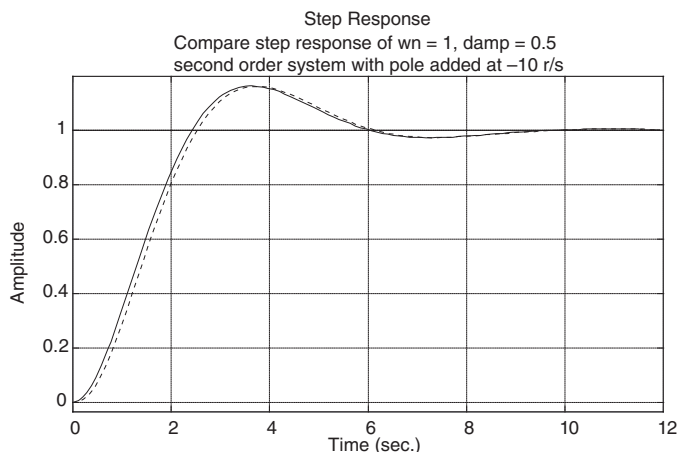
¹¹ For more detail, see e.g., James Roberge, *Operational Amplifiers: Theory and Practice*, John Wiley, 1975.

Second-order electrical system with extra high-frequency poles

How does a second-order system behave if there are one or more extra poles added? Well, the response depends on how close to the resonant poles the extra added pole(s) are. As we have seen previously, a high-frequency pole behaves as a magnitude of approximately 1.0 with a negative phase shift that decreases approximately proportionally to frequency (corresponding to a time delay).

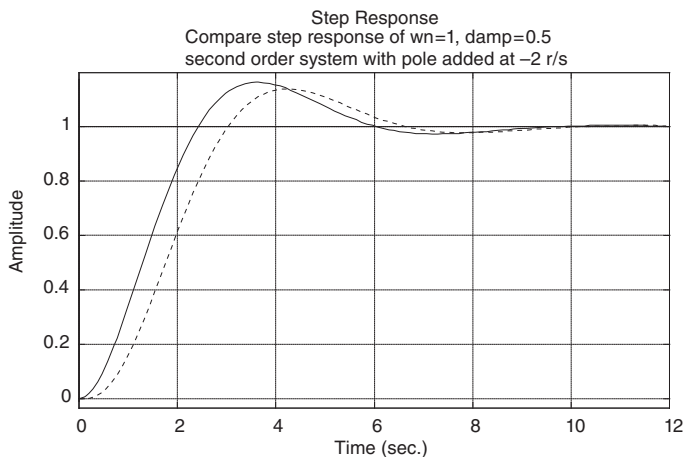
Shown in **Figure 2-19** is the step response for a second-order system with natural frequency $\omega_n = 1$ radian/second and damping ratio $\zeta = 0.5$. Superimposed on this plot is a second plot showing the effects of an additional pole at -10 radians/second. Note that the two plots match up fairly well, with the exception that the plot for the system with the extra pole is delayed (by approximately 0.1 second) in the risetime area. This corresponds to the finite delay caused by the high-frequency pole.

Figure 2-19: Step responses for second-order systems with natural frequency $\omega_n = 1$ radian/s and damping ratio $\zeta = 0.5$, with and without extra pole at -10 radian/s.



In **Figure 2-20**, we see the effects of a lower frequency pole added to the same system. This time, the extra pole is at -2 radians/second. We note that the effect on the original system is more pronounced this time. This is expected, since the added pole isn't that much higher than the natural frequency of the original system.

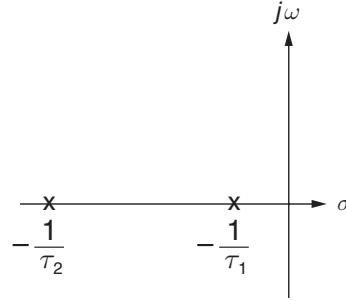
Figure 2-20: Step response for second-order systems with $\omega_n = 1$ radian/s and damping ratio $\zeta = 0.5$, with extra pole at -2 radian/s. Solid line: original second-order system. Dashed line: second-order system with extra higher-frequency pole added.



Second-order system with widely spaced real-axis poles

A second-order system with widely spaced real-axis poles is shown in **Figure 2-21**.

Figure 2-21: Second-order system with two widely spaced real-axis poles.



The transfer function for this system is:

$$H(s) = \frac{1}{(\tau_1 s + 1)(\tau_2 s + 1)} \quad [2-44]$$

where τ_1 is the time constant associated with the dominant, low-frequency pole, and τ_2 is the time constant of the faster (nondominant) pole. Multiplying out the denominator results in:

$$H(s) = \frac{1}{\tau_1 \tau_2 s^2 + (\tau_1 + \tau_2)s + 1} \quad [2-45]$$

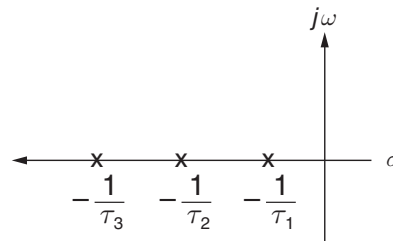
In this case, the second-order system behaves as two cascaded first-order systems.

Finding approximate pole locations from the transfer function denominator

We'll now go off on a little bit of a tangent which will help us in further chapters when we'll be approximating the bandwidth of circuits. Let's assume that we have a system, and that we know that this system has three poles that are on the real axis and are widely spaced, as shown in **Figure 2-22**. We have a low-frequency pole at frequency $-1/\tau_1$, a high-frequency pole at $-1/\tau_3$, and an intermediate frequency pole at $-1/\tau_2$, or, said another way, the pole time constants are related as $\tau_1 \gg \tau_2 \gg \tau_3$. The transfer function of this system is:

$$H(s) = \frac{1}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1)} = \frac{1}{\tau_1 \tau_2 \tau_3 s^3 + (\tau_1 \tau_2 + \tau_1 \tau_3 + \tau_2 \tau_3)s^2 + (\tau_1 + \tau_2 + \tau_3)s + 1} \quad [2-46]$$

Figure 2-22: Pole plot of three widely spaced poles on the negative real axis.



This transfer function is of the form:

$$H(s) = \frac{1}{a_3 s^3 + a_2 s^2 + a_1 s + 1}$$

$$a_3 = \tau_1 \tau_2 \tau_3$$

$$a_2 = \tau_1 \tau_2 + \tau_1 \tau_3 + \tau_2 \tau_3$$

$$a_1 = \tau_1 + \tau_2 + \tau_3$$
[2-47]

Let's call the pole locations p_{low} , p_{medium} and p_{high} for “low,” “medium,” and “high” frequency. We find that the approximate pole locations are:

$$p_{low} \approx -\frac{1}{a_1} \approx -\frac{1}{\tau_1 + \tau_2 + \tau_3} \approx -\frac{1}{\tau_1}$$

$$p_{high} \approx -\frac{a_2}{a_3} \approx -\frac{\tau_1 \tau_2 + \tau_1 \tau_3 + \tau_2 \tau_3}{\tau_1 \tau_2 \tau_3} \approx -\frac{\tau_1 \tau_2}{\tau_1 \tau_2 \tau_3} \approx -\frac{1}{\tau_3}$$

$$p_{medium} \approx -\frac{a_1}{a_2} \approx -\frac{\tau_1 + \tau_2 + \tau_3}{\tau_1 \tau_2 + \tau_1 \tau_3 + \tau_2 \tau_3} \approx -\frac{\tau_1}{\tau_1 \tau_2} \approx -\frac{1}{\tau_2}$$
[2-48]

In fact, we can extend this methodology to denominator polynomials of any degree where the pole locations are all on the negative real axis and are widely spaced. In this case, the pole locations are:

$$p_k \approx -\frac{a_{k-1}}{a_k}$$

$$a_o = 1$$
[2-49]

As an example, consider the fifth-order polynomial

$$H(s) = \frac{1}{10^{-10} s^5 + 1.111 \times 10^{-6} s^4 + 1.122 \times 10^{-3} s^3 + 1.122 \times 10^{-1} s^2 + 1.111 s + 1}$$
[2-50]

Using our approximation formula, we find the approximate pole locations:

$$p_5 \approx -\frac{1.111 \times 10^{-6}}{10^{-10}} \approx -11,110 \text{ r/s}$$

$$p_4 \approx -\frac{1.122 \times 10^{-3}}{1.111 \times 10^{-6}} \approx -1,010 \text{ r/s}$$

$$p_3 \approx -\frac{1.122 \times 10^{-1}}{1.122 \times 10^{-3}} \approx -100 \text{ r/s}$$

$$p_2 \approx -\frac{1.111}{1.122 \times 10^{-1}} \approx -9.9 \text{ r/s}$$

$$p_1 \approx -\frac{1}{1.111} \approx -0.9 \text{ r/s}$$
[2-51]

In this example, the actual pole locations¹² are at -1 , -10 , -100 , -1000 and -10000 radians per second. We find that the methodology gives us a rough approximation of the pole locations if we're correct in our initial assumption that the poles are on the real axis and are widely spaced.

Review of Resonant Electrical Circuits

Consider a parallel LC circuit as shown in **Figure 2-23**. This is a standard resonant circuit that will oscillate at frequency ω_r as the energy sloshes back and forth between electric storage in the capacitor and magnetic storage in the inductor. Let's assume that the initial condition in the circuit is that the inductor current is zero and the capacitor voltage is some voltage v_c . The current in the inductor is given by the differential equation:

$$\frac{di_L}{dt} = \frac{v_c}{L} \quad [2-52]$$

Similarly, the voltage on the capacitor can be found by:

$$\frac{dv_c}{dt} = \frac{-i_L}{C} \quad [2-53]$$

We can solve for the voltage on the capacitor $v(t)$ by differentiating the capacitor equation and substituting for di/dt , resulting in:

$$\frac{d^2 v_c}{dt^2} = -\frac{1}{C} \frac{di_L}{dt} = -\frac{v_c}{LC} \quad [2-54]$$

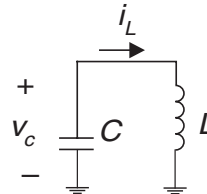
Now, we can find the resonant frequency by guessing that the voltage $v(t)$ is sinusoidal with $v(t) = V_o \sin \omega t$. Putting this into the equation for capacitor voltage results in:

$$-\omega^2 \sin(\omega t) = -\frac{1}{LC} \sin(\omega t) \quad [2-55]$$

This means that the resonant frequency is the standard (as expected) resonance:

$$\omega_r^2 = \frac{1}{LC} \quad [2-56]$$

Figure 2-23: Undamped parallel LC resonant circuit.



¹² If you don't believe me, multiply it out for yourself or solve it using MATLAB.

Use of Energy Methods to Analyze Undamped Resonant Circuits

The use of energy methods can easily solve many problems in electrical and mechanical engineering. Energy is stored in capacitors in an electric field, in inductors in a magnetic field, in moving objects, springs and thermal masses. The stored energies in these various storage modes are found in **Table 2-1**.

Table 2-1: List of symbols.

Storage Mode	Relationship	Comments
Capacitor/electric field storage	$E_{elec} = \frac{1}{2}CV^2$	C = capacitance. V = voltage.
Inductor/magnetic field storage	$E_{mag} = \frac{1}{2}LI^2 = \int \frac{B^2}{2\mu_o} dV$	L = inductance; I = current; B = magnetic flux density (Tesla).
Kinetic energy	$E_k = \frac{1}{2}Mv^2$	M = mass, v = velocity.
Rotary energy	$E_r = \frac{1}{2}I\omega^2$	$I \equiv$ mass moment of inertia (kg- m ²). ω is rotary speed in radians/second
Spring	$E_{spring} = \frac{1}{2}kx^2$	$k \equiv$ spring constant (N/m). x = spring displacement.
Potential energy	$\Delta E_p = Mg\Delta h$	$\Delta h \equiv$ height change
Thermal energy	$\Delta E_T = C_{TH}\Delta T$	$C_{TH} \equiv$ thermal capacitance (J/K). ΔT is change in temperature (K)

By using energy methods we can find the ratio of maximum capacitor voltage to maximum inductor current in the resonant circuit of **Figure 2-23**. Assuming that the capacitor is initially charged to V_o volts, and remembering that capacitor stored energy $E_c = \frac{1}{2}CV^2$ and inductor stored energy is $E_L = \frac{1}{2}LI^2$, we can write the following:

$$\frac{1}{2}CV_o^2 = \frac{1}{2}LI_o^2 \quad [2-57]$$

What does this mean about the magnitude of the inductor current? Well, we can solve for the ratio of V_o/I_o resulting in:

$$\frac{V_o}{I_o} = \sqrt{\frac{L}{C}} \equiv Z_o \quad [2-58]$$

The term “ Z_o ” is defined as the characteristic impedance of a resonant circuit. Let’s assume that we have an inductor-capacitor circuit with $C = 1$ microfarad and $L = 1$ microhenries. This

means that the resonant frequency is 10^6 radians/second (or 166.7 kHz) and that the characteristic impedance is 1 ohm. Shown below is a simulation of this circuit with an initial condition on the capacitor of 1 volt. As expected, the resonant frequency is 166 kHz and the maximum inductor current is 1 amp. Note that the capacitor voltage and inductor current are 90 degrees out of phase. This is characteristic of an undamped resonant circuit operating at resonance.

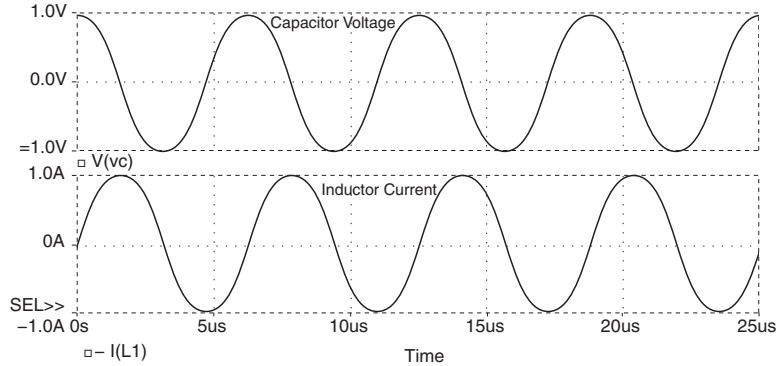


Figure 2-24: PSPICE output for parallel undamped LC circuit with $L = 1 \mu\text{H}$ and $C = 1 \mu\text{F}$ and initial conditions with the capacitor charged to 1 volt, and zero inductor current. The resultant oscillation frequency is $\sim 167 \text{ kHz}$.

We could also use energy methods to analyze the oscillation frequency of the simple mass-spring system evaluated before. The maximum potential energy stored in a spring is:

$$E_p = \frac{1}{2} k (Y_o)^2 \quad [2-59]$$

The maximum kinetic energy stored in the moving mass is:

$$E_k = \frac{1}{2} M v^2 = \frac{1}{2} M (\omega Y_o)^2 \quad [2-60]$$

If we equate maximum kinetic and potential energy, the result is:

$$\frac{1}{2} k (Y_o)^2 = \frac{1}{2} M (\omega Y_o)^2 \rightarrow \omega = \sqrt{\frac{k}{M}} \quad [2-61]$$

Transfer Functions, Pole/Zero Plots and Bode Plots

Shown in **Table 2-2** are numerous transfer functions, their corresponding pole/zero plots and Bode frequency response plots.

Table 2-2: Transfer functions with their pole/zero and Bode frequency response plots.

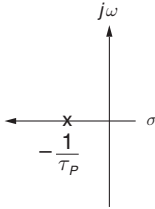
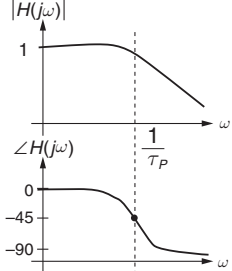
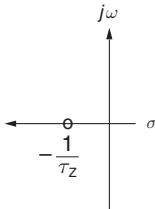
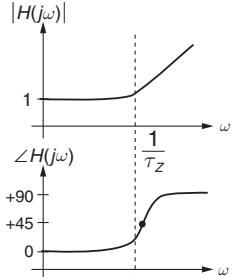
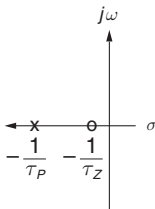
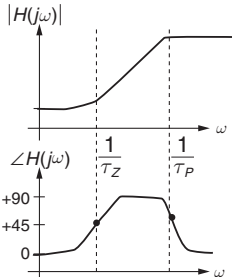
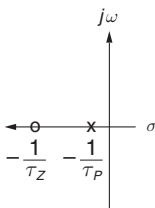
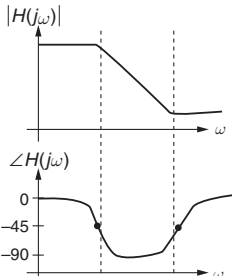
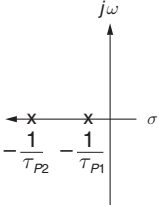
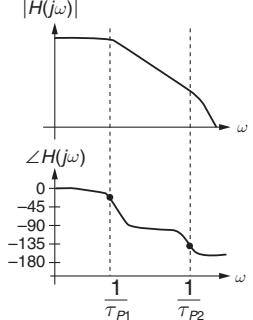
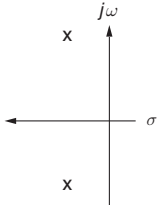
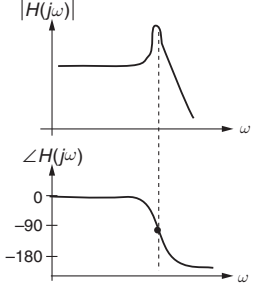
System Type	Transfer Function $H(s)$	Pole/Zero Plot	Bode Plot
Single Pole	$\frac{1}{\tau_p s + 1}$		
Zero	$\tau_z s + 1$		
Lead	$\frac{\tau_z s + 1}{\tau_p s + 1}, (\tau_z > \tau_p)$		
Lag	$\frac{\tau_z s + 1}{\tau_p s + 1}, (\tau_z < \tau_p)$		

Table 2-2: Transfer functions with their pole/zero and Bode frequency response plots (continued).

System Type	Transfer Function $H(s)$	Pole/Zero Plot	Bode Plot
Two Pole	$\left(\frac{1}{\tau_{p1}s+1}\right)\left(\frac{1}{\tau_{p2}s+1}\right)$		
Second order, low damping	$\frac{1}{\left(\frac{s}{\omega_n}\right)^2 + 2\zeta\left(\frac{s}{\omega_n}\right) + 1}$		

Risetime for Cascaded Systems

For multiple systems in cascade (**Figure 2-25**), we can figure out the overall system risetime if we know the risetime of the individual blocks. The risetimes do not simply add; for instance, for N systems wired in series, each with its own risetime $\tau_{R1}, \tau_{R2}, \dots, \tau_{RN}$, the overall risetime of the cascade τ_R is:¹³

$$\tau_R = \sqrt{\tau_{R1}^2 + \tau_{R2}^2 + \tau_{R3}^2 + \dots + \tau_{RN}^2} \quad [2-62]$$

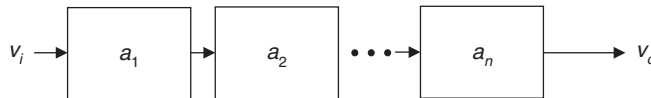


Figure 2-25: Cascaded systems.

Using the risetime addition rule, we find that a cascade of N identical stages has a risetime equal to \sqrt{N} times the risetime of an individual stage.

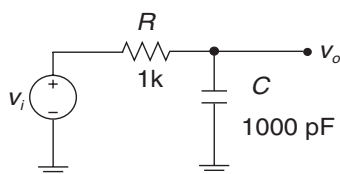
¹³ A detailed mathematical treatment for risetimes of cascaded systems can be found in William Siebert's *Circuits, Signals and Systems*, chapter 16, and also in Tom Lee's *The Design of CMOS Radio-Frequency Integrated Circuits*, chapter 7. Note that the risetime addition rule only holds true if the step responses of the individual systems are well-behaved.

Chapter 2 Problems

Problem 2.1

A single-pole low-pass filter (**Figure 2-26**) has $R = 1\text{ k}\Omega$ and $C = 1000\text{ pF}$. Find the input-output transfer function $H(s)$ and calculate the 10–90% risetime and -3dB bandwidth of this circuit.

Figure 2-26: Single-pole filter



Problem 2.2

Find and plot the angle and group delay of the single pole filter of Problem 2.1.

Problem 2.3

Simulate the circuit of Problem 2.1 and verify your results in Problems 2.1 and 2.2.

Problem 2.4

Design a second-order RLC circuit with an input-output transfer function exhibiting a natural frequency $\omega_n = 10^6$ radians/second and a damping ratio $\zeta = 0.1$. What is the Q of this circuit ?

Problem 2.5

For the circuit of Problem 2.4, plot the magnitude and angle of the transfer function. Correlate the simulated peak in the magnitude response with circuit Q .

Problem 2.6

A mass-spring system has mass $M = 10\text{ kg}$ and spring constant $k = 10\text{ Newtons/meter}$. What is the oscillation frequency of this system?

Problem 2.7

An undamped parallel resonant circuit has $L = 10\text{ microhenries}$ and $C = 100\text{ microfarads}$. At time $t = 0$, the capacitor is charged to 1 V and the inductor current is zero. Find and plot capacitor voltage and inductor current for $t > 0$. What is the peak inductor current?

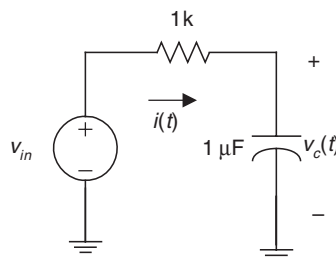
Problem 2.8

In the RC circuit of **Figure 2-27**, the circuit is driven by a 100-volt input step $v_{in}(t) = 100u_{-1}(t)$.

- Sketch and label resistor current $i(t)$ and capacitor voltage $v_c(t)$.
- How much energy is dissipated in the resistor during the capacitor charging process?
- Increase the resistor to $10\text{ k}\Omega$ and repeat the experiment. After you apply the step and the capacitor is fully charged, how much energy is dissipated in the resistor?

(Hint: You can do (b) and (c) the easy way or the hard way. The easy way considers the energy stored and/or charge stored in the capacitor.)

Figure 2-27: Circuit for Problem 2.8.



The circuit with $R = 10\text{ k}$ is now driven by a 100-volt square wave at 10 Hz (for $v_{in}(t)$, $V_{max} = 100\text{ V}$, $V_{min} = 0\text{ V}$).

- Sketch and label resistor current $i(t)$ and capacitor voltage $v_c(t)$.
- Approximately how much power is dissipated in the resistor?
- Increase the frequency to 1 MHz. After many RC time constants, approximately how much power is dissipated in the resistor? (Hint: What frequency components are in a 50% duty-cycle square wave?)

Now, the circuit with $R = 10\text{ k}$ is driven with an AC voltage source $v_{in}(s)$.

- Draw the Bode plot of v_o/v_{in} .
- What is the -3 dB bandwidth, in Hz of v_o/v_{in} ?
- What is the 10–90% risetime?
- Derive an expression relating the bandwidth and risetime.

Problem 2.9

For the LC circuit in **Figure 2-28**,

- Calculate the resonant frequency.
- Sketch the step response of the capacitor voltage, labeling values and axes.

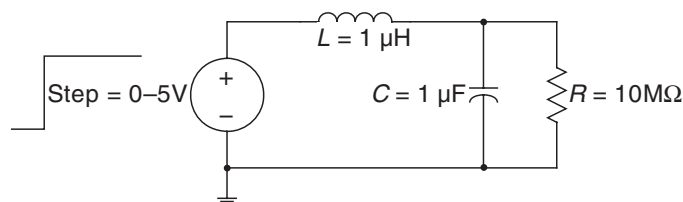


Figure 2-28: Circuit for Problem 2.9.

Problem 2.10

- Find the 10–90% risetime of the low-pass filter of **Figure 2-29a**.
- The low-pass filter circuit from part (a) is cascaded with an identical circuit, with a buffer in-between to ensure that one stage doesn't load down the other. Find the 10–90% risetime of the overall cascade.
- Verify these results using PSPICE.

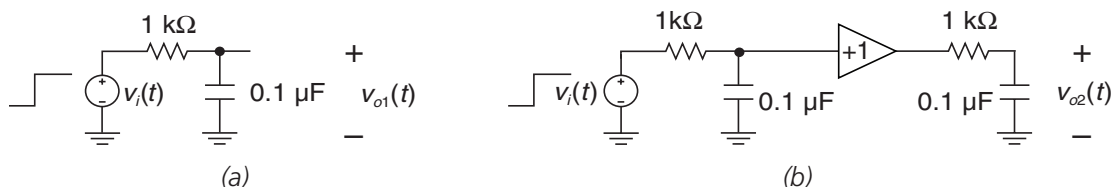


Figure 2-29: (a) Single-pole system.

(b) Two cascaded first-order systems with unity gain buffer in between.

Problem 2.11

You have an RC circuit configured as a one-pole low-pass filter and are attempting to ascertain the capacitance value of the circuit by measuring the step response. You know that the resistance value is 100Ω . All you have available is an oscilloscope with a bandwidth of 100 MHz and a step generator with a very fast risetime. Using the oscilloscope, you measure a 10–90% risetime of 5 ns. What is the capacitance value?

Problem 2.12

A roller coaster is initially at rest at the top of an 84-foot (25.6-meter) tall hill (**Figure 2-30a**). The coaster is then gently pushed over the edge and travels to the bottom of the hill as shown in **Figure 2-30b**. What is the speed of the coaster at the bottom of the hill, assuming that mechanical and air friction is negligible?

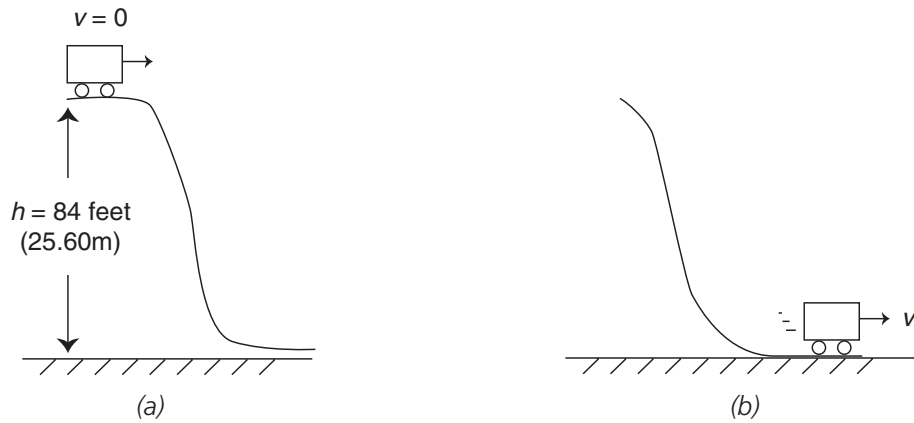


Figure 2-30: Energy method problem. (a) Roller coaster on top of 84-foot high hill.
(b) Roller coaster at bottom of hill.

References

- Andrews, James, “Low-Pass Risettime Filters for Time Domain Applications,” *Picosecond Pulse Labs*, Application note #AN-7a, 1999.
- Beranek, Leo L., *Acoustics*, Acoustical Society of America, 1954.
- CRC Press, *CRC Standard Mathematical Tables*, 28th edition, 1987.
- Guillemin, Ernst, *Introductory Circuit Theory*, John Wiley, 1953.
- Johnson, Howard, “Risettime of Lossy Transmission Lines,” *EDN*, October 2, 2003, p. 32.
- Lee, Thomas H., *The Design of CMOS Radio-Frequency Integrated Circuits*.
- Rao, Singiresu S., *Mechanical Vibrations*, 3rd edition, Addison-Wesley, 1995.
- Roberge, James, *Operational Amplifiers: Theory and Practice*, John Wiley, 1975.
- Senturia, Stephen D., and Wedlock, Bruce D., *Electronic Circuits and Applications*, reprinted by Krieger, 1993.
- Siebert, William McC., *Circuits, Signals and Systems*, McGraw-Hill, 1986.

Review of Diode Physics and the Ideal (and Later, Nonideal) Diode

In This Chapter

- *The basics of bipolar devices are covered, including basic semiconductor physics,¹ the concepts of electron and hole flow in semiconductors, the differences between drift and diffusion flow, generation and recombination, and the effects of semiconductor doping on carrier concentrations. We finish with a discussion of the ideal diode, and illustrate how a diode can conduct forward current, but can also block reverse voltage. Detailed mathematical derivations are avoided wherever possible. However, enough mathematical detail is given so that the reader can discern the important scaling laws and functional dependencies of the ideal diode. At the end of the chapter we'll discuss some of the factors that result in nonideal behavior in diodes. We'll conclude with a discussion of load lines, a useful method for solving for the operating point of circuits with nonlinear devices. The load-line technique will be useful in later chapters in analyzing transistors.*

Current Flow in Insulators, Good Conductors and Semiconductors

In nature, from the point of view of the ease of producing current flow in a material, there are three broad classes of materials: insulators, conductors and semiconductors. Semiconductors and metals can support significant current flow but the charge movement mechanisms are different in the two types of materials. How “good” an electrical conductor is can be quantified by material property electrical resistivity and/or its inverse, electrical conductivity. Electrical resistivity² is a measure of how well a given material conducts current. If there are lots of free charged carriers available, a material is deemed a good conductor.

¹ We will not go into the quantum mechanics of semiconductors, which provides the rigorous analyses. The simpler models developed in this chapter hopefully will give insight into the basics of semiconductor operation. Excellent reviews of semiconductor physics are given in Shockley's and Bardeen's 1956 Nobel prize lectures, with reference given at the end of this chapter.

² If you look in technical references (such as the *Handbook of Chemistry and Physics*) you can often find electrical conductivity listed for materials. Electrical resistivity has units of ohms-meters. Electrical conductivity is the inverse of electrical resistivity.

Insulators are materials such as quartz, rubber, plastics and certain ceramics which do not support current flow very well. In other words, the electrical resistivity of an insulating material is very high. The resistance to current flow is very high in insulators because there are very few free charges available to contribute to current flow.

Good conductors support current flow easily. In other words, to get substantial current flow in a good conductor you don't have to supply very much driving voltage. Some metals, such as copper, aluminum, gold and silver, are very good conductors because they have a sea of free electrons, each with a negative charge $-q$ associated with them,³ available to support current flow. Typical metals have electrical conductivity many orders of magnitude higher than semiconductors or insulators. Most metals are good conductors; however, none of them are perfect conductors,⁴ at least at room temperature.

Let's do a simple calculation of the resistance of a copper rod (**Figure 3-1**). The electrical resistance of the rod⁵ is:

$$R = \frac{l}{\sigma A} \quad [3-1]$$

where l is the length, A is cross-sectional area and σ is electrical conductivity (which has units of $\Omega^{-1}\text{m}^{-1}$). Let's connect a battery to the metal rod as shown in **Figure 3-1**. The current flow is left to right in the metal rod (corresponding to electron flow from right to left⁶). For a metal rod of length 0.1 meters and cross-sectional area 10^{-4} m^2 , the resistance at room temperature⁷ is approximately 16.9 microohms.

³ The charge on an electron is $-q$, where $q \approx 1.6 \times 10^{-19}$ Coulombs.

⁴ There are compounds, such as niobium-titanium, which become superconducting at cryogenic temperatures (e.g., liquid helium temperature, at 4.2K). Scientists are searching for superconductors (conductors whose electrical resistivity goes to zero if the temperature is lowered below a critical temperature). As of yet, no room-temperature superconductors have been discovered. The electrical conductivity of copper at room temperature is approximately $5.9 \times 10^7 \Omega^{-1}\text{m}^{-1}$, while that of aluminum is approximately $3.5 \times 10^7 \Omega^{-1}\text{m}^{-1}$. Note that these values of electrical conductivity vary significantly with the specific alloy, and that electrical conductivity of metals increases as operating temperature decreases.

⁵ This calculation ignores any resistance due to contacts at the ends of the rod.

⁶ Note that the charge on the electron is negative. Hence, when electrons flow from right to left, the net current flow is from left to right. The directions of current flow are shown by convention.

⁷ Electrical resistivity is the inverse of electrical conductivity. The electrical resistivity of copper at room temperature is approximately $1.7 \times 10^{-8} \Omega\text{-m}$. The temperature coefficient of resistivity of copper is approximately +0.4% per degree C.

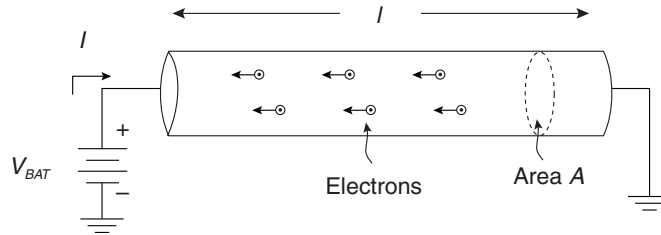


Figure 3-1: Metal rod with current flow due to applied battery voltage. In the metal, electric current is carried in a flow of electrons. Electrons flowing from right to left carry a negative charge, and hence the net current flow is from left to right.

In terms of electrical resistivity, somewhere in-between conductors and insulators are semiconductors. These materials are moderately good at supporting current flow. Typical semiconductors are silicon (Si), germanium (Ge), and gallium arsenide (GaAs). An important difference between semiconductors and metallic conductors is the method by which charged carriers convey current. In semiconductors, two types of charged carriers contribute to current flow. They are electrons (negatively charged particles with charge $-q$) and holes (positively charged particles with charge $+q$). An *intrinsic* semiconductor is a semiconductor material that has not been doped with impurities. In the next section, we'll discuss electrons and holes in semiconductors.

Electrons and Holes

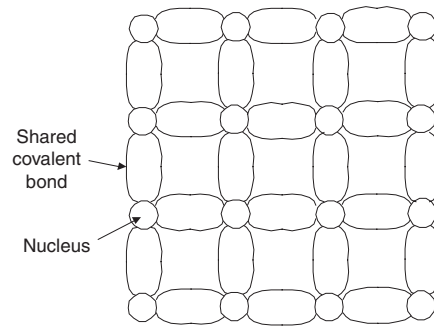
Shown in an idealized two-dimensional way in **Figure 3-2** is a silicon lattice.⁸ Each silicon atom has four outer shell (“valence band”) electrons and four shared outer electrons with the nearest atomic neighbors. Hence there are eight total “covalent” bonds per atom. This idealized model of the silicon lattice and the resultant concepts of electron and hole motion gloss over many subtleties associated with quantum mechanical effects. However, this idealized model has proven useful, and surprisingly accurate, in explaining current flow in diodes and transistors.

⁸ The structure of the silicon and germanium lattice is actually a diamond-type crystal lattice with atoms having covalent bonds to four nearest neighbors. This picture is an idealized, flattened-out depiction of the actual lattice. This method of explaining approximately carrier flow in semiconductors was put forth by physicist William Shockley in the early days of semiconductors (about 1950). The laws of physics that explain electronic bonding are quantum mechanical in nature, and hence Schrodinger's wave equation must be solved. Says Shockley in describing carrier motion in transistors, “The laws of physics which explain the behavior of the electron-pair bond are *quantum* laws and employ *wave-mechanics* to describe the motion of the electrons.” See W. Shockley “Transistor Electronics: Imperfections, Unipolar and Analog Transistors.” Hence the models shown in **Figure 3-2** and **Figure 3-3** are a bit cartoonish in nature; however, they capture the essence of the physics involved in carrier flow in semiconductors.

Chapter 3

In the ideal lattice, all of the covalent bonds are filled and the associated electrons are tightly bound to their neighboring nuclei. Therefore, no “free” charges are available to contribute to current flow.⁹

Figure 3-2: Structure of idealized silicon lattice with all covalent bonds filled.

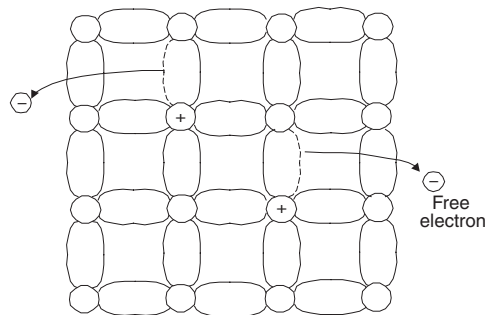


Now, in a real-world semiconductor, some fraction of these covalent bonds break due to the random thermal energy of the silicon lattice (**Figure 3-3**). If a voltage is applied to the sample, these free electrons can flow and support a net current. The resultant space where the electron was before the bond was broken is called a *hole*. A hole carries a net positive charge of $+q$, and holes can *also* move about in the silicon lattice. In other words, the hole behaves like a charge carrier much like the electron, except that the hole carries a positive charge. The total current flow in a semiconductor is hence a combination of the current due to free electrons and the current due to free holes, or:

$$J = J_e + J_h \quad [3-2]$$

where J is total current density (expressed as amps per square meter¹⁰) and J_e and J_h are the electron and hole currents, respectively.

Figure 3-3: Real-world silicon lattice with some broken covalent bonds. Each broken bond allows a negatively charged electron and a positively charged hole to move.



⁹ This is how a semiconductor behaves at very low temperatures. At cryogenic temperatures, for example, few outer shell bonds are broken (since the lattice is so cold and there is not enough thermal energy to shake electrons free).

¹⁰ It's simple to find the total current in a semiconductor if you can calculate the current density J . Just multiply current density J (with units of amps/m²) by the total cross-sectional area A through which the current flows: $I = JA$.

In a perfect silicon lattice without any added impurities, which we'll call an *intrinsic semiconductor*, the number of free electrons and the number of free holes are equal, because free electrons and holes are created in pairs. We'll call the number of electrons per unit volume in an intrinsic semiconductor n , and the number of holes per unit volume p . From this simple reasoning, we find that the number of holes equals the number of electrons in an intrinsic semiconductor, or:

$$n = p \equiv n_i \quad [3-3]$$

We have now defined the *intrinsic carrier concentration* n_i , which is the number of bonds that are broken in the intrinsic semiconductor.

Furthermore, we also find that, as the temperature rises, the number of broken bonds increases since there is more thermal agitation. Therefore, in the intrinsic semiconductor, the intrinsic carrier concentration (we'll call it $n_i(T)$) is a function of temperature,¹¹ or:

$$np = n_i^2(T) \quad [3-4]$$

In silicon, a typical value for intrinsic carrier concentration n_i at room temperature is $\sim 1.5 \times 10^{10}$ per cubic centimeter. This intrinsic carrier concentration is a strong function of temperature and increases with temperature.

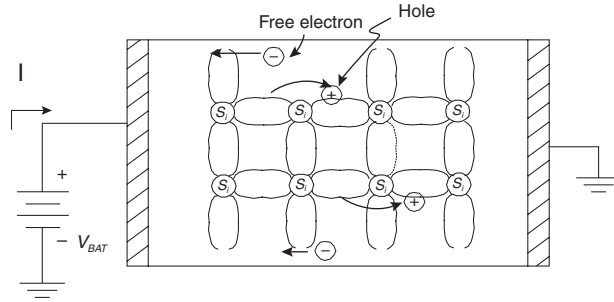
Let's do a simple thought experiment, and connect a battery to a piece of semiconductor material, in this case, silicon (**Figure 3-4**). The intrinsic silicon has some fraction¹² of its total outer-shell bonds broken and hence there is concentration n_i of free electrons and free holes in the sample. These free charges can move and contribute to conduction current. The electric field set up by the battery causes the free electrons to flow to the left, and causes holes to flow to the right.¹³ This total current adds up to the battery terminal current I as shown.

¹¹ This result is called the "mass-action" law, and has a counterpart in chemistry. For instance, in an acidic solution, the concentrations of hydrogen (H^+) and hydroxyl (OH^-) also follow the mass action law.

¹² Let's take a look at the fraction of total bonds that are broken. Silicon contains about 5×10^{22} atoms per cubic centimeter of volume. At room temperature, there are about $n_i \approx 1.5 \times 10^{10}$ free electrons and free holes per cubic centimeter. So, a *very* small fraction (approximately one part in 3.33×10^{12}) of total electron bonds are broken. However, this small fraction has a drastic effect on current flow in semiconductors.

¹³ We can think of a hole as a void that gets filled when a covalent bond jumps from one location to another. In this simplified way we can think of a "hole" as a current-carrying particle.

Figure 3-4: Battery connected to a piece of semiconductor material. The semiconductor has current flow both due to negatively charged electrons and positively charged holes. In the picture shown, electrons move from right to left, and holes move from left to right, due to the force on the charges from the applied electric field from the battery.



Drift, Diffusion, Recombination and Generation

When a charge carrier is free in a semiconductor, several things can happen to it. It can *drift* due to a force applied to it from an electric field. It can *diffuse* if there is a density gradient. If an electron meets a hole, the electron and hole can *recombine* wherein the electron fills the hole. All the while this is happening, new carriers are being *generated* due to thermal agitation or external excitation.¹⁴ All four mechanisms happen simultaneously in a semiconductor. The details of the balance of drift, diffusion, recombination and generation tell us the net current in a piece of semiconductor.

Drift

Shown in **Figure 3-5a** is an electron in an applied electric field \mathbf{E} . The electron has an electrostatic force acting on it. In this picture, the force acts from right to left if the electric field vector \mathbf{E} points from left to right. The electron has an average velocity given by:

$$v = -\mu_n \mathbf{E} \quad [3-5]$$

where μ_n is a material constant¹⁵ called the *mobility* of the electron.¹⁶ The velocity has a minus sign because the drift motion of the electron is opposite of the direction of the applied electric field. Shown in **Figure 3-5a** is a hole in an applied electric field \mathbf{E} . The hole has a force acting on it in the same direction as the applied field. The velocity of the hole is given by:

$$v = \mu_p \mathbf{E} \quad [3-6]$$

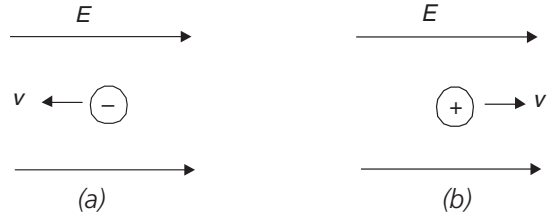
¹⁴ For instance, if you shine light on a piece of semiconductor or heat it up you will generate electron/hole pairs.

¹⁵ Note that this “constant” is not really constant at all. It varies somewhat with temperature, doping level, and level of the applied electric field. You might ask why the electron has an average velocity v and why it doesn’t keep accelerating to higher and higher velocity. A perfectly periodic lattice at absolute zero would not scatter a free electron. However, above absolute zero the lattice vibrates and the moving electron bangs into the lattice, slowing it down. The moving electrons can also be scattered by impurities in the lattice. The average time between collisions we’ll call τ_c . The electron reaches an average drift velocity proportional to the electric field. The mobility of a free particle is given by: $\mu = q\tau_c/m_{eff}$ where q is electronic charge and m_{eff} is the effective mass of the moving particle.

¹⁶ Mobility has units of square centimeters per volt-second.

where μ_p is the mobility of the hole. Note that the mobilities of electrons and holes are different. Typically, the mobility of electrons is two to three times higher than holes in semiconductors.¹⁷

Figure 3-5:
Charged carriers in an electric field.
(a) Electron drift, in a direction opposite to the electric field.
(b) Hole drift, in the same direction as the electric field.



Now, let's consider the case where we have a lot of holes in the presence of an electric field. The drift current density due to holes is:

$$J_{h,drift} = q\mu_p pE \quad [3-7]$$

where p is the concentration of holes (holes per unit volume) and E is the electric field. Similarly, drift current density due to electrons is:

$$J_{n,drift} = q\mu_n nE \quad [3-8]$$

Diffusion

Fick's law of diffusion describes how particles under random thermal motion tend to diffuse¹⁸ from a region of higher concentration to a region of lower concentration. This principle is illustrated by opening a perfume bottle in the corner of a closed room. If you wait long enough, the perfume odor will permeate the room because the perfume molecules have diffused from one side of the room to the other, from a region of high concentration to a region of low concentration. Mathematically, three-dimensional diffusion is characterized by Fick's diffusion law, which states that the diffusion flux is proportional to the concentration gradient, as:

$$\mathbf{F} = -D\nabla C \quad [3-9]$$

where C is the concentration of the diffusing particles, F is the diffusion flux (particles per square meter per second) and D is the diffusion constant which has units of $\text{cm}^2/\text{second}$. For a one-dimensional problem, Fick's law reduces to:

$$\mathbf{F} = -D \frac{dC}{dx} \quad [3-10]$$

Therefore, charged particles tend to flow down a concentration gradient. This diffusion process also occurs in PN junctions whenever there are gradients of free charged carriers.

We can work out the form (but not the detail) of Fick's law by considering a thought experiment. Consider a region of space where there is a changing concentration of free charges, in

¹⁷ In silicon at modest doping levels and electric fields, typical numbers for mobilities are $\mu_n \approx 1360 \text{ cm}^2/\text{V-sec}$ and $\mu_p \approx 500 \text{ cm}^2/\text{V-sec}$. See, e.g., R. W. Pierret, *Modular Series on Solid State Devices*, vol. 1, "Semiconductor Fundamentals," p. 58.

¹⁸ A good, intuitive explanation of diffusion is given in R. W. Pierret, *Modular Series on Solid State Devices*, vol. 1, "Semiconductor Fundamentals," pp. 67–71 and also in R. Feynman, et. al, *The Feynman Lectures on Physics*, vol. 1, chapter 43.

this case, holes (**Figure 3-6**). The holes are undergoing random thermal motion. For instance, at $x = -x_o$, on average, one half of the holes are traveling to the left, and one half are traveling to the right. The same is true at $x = +x_o$. In order to find the net current at $x = 0$, we recognize that the current at $x = 0$ is the sum of the current from the left plus the current from the right, or:

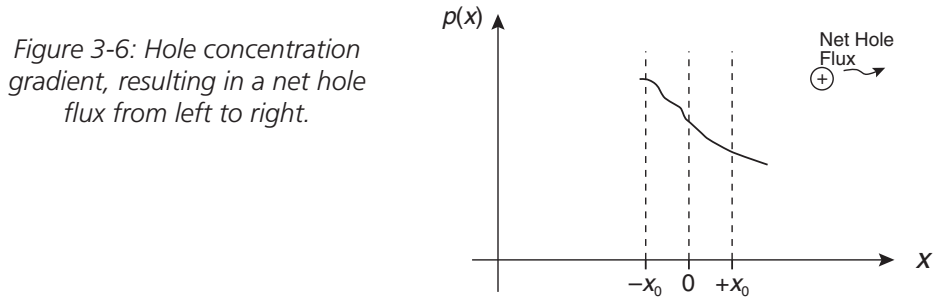
$$J(x=0) = k[p(x=-x_o) - p(x=+x_o)] = -k(2x_o) \left[\frac{dp}{dx} \right] \quad [3-11]$$

where k is some constant that makes the units work out. Note that the net current at $x = 0$ is proportional to the difference in the concentrations at $x = -x_o$ and $x = +x_o$. Through a mathematical manipulation¹⁹ above we see that the current is also proportional to dp/dx , or the gradient of the concentration. We now recognize the familiar form for diffusion current for holes:

$$J_{h,diff} = -qD_p \frac{dp}{dx} \quad [3-12]$$

where q is the electronic charge and D_p is the diffusion constant²⁰ for holes. Using a similar derivation, we can find the electron diffusion current:

$$J_{e,diff} = qD_n \frac{dn}{dx} \quad [3-13]$$



Let's do a diffusion thought experiment, illustrated in **Figure 3-7**. At time $t = 0$, a high concentration of particles (in this case, electrons) exists at $x = 0$. These particles can be created by illuminating a piece of semiconductor, or by other mechanisms. The particles are in random thermal motion; some diffuse to the left, and some diffuse to the right. The concentration of particles $n(x,t)$ at various times is shown. At $t = t_1$, the maximum concentration at

¹⁹ This hand-waving derivation glosses over details of collisions and mean-free-path, etc., which is well discussed in Feynman's lecture 43; see reference at the end of this chapter.

²⁰ Approximate values for diffusion constants in silicon at room temperature are $D_n \approx 35 \text{ cm}^2/\text{sec}$ and $D_p \approx 12.5 \text{ cm}^2/\text{sec}$. Interestingly enough, the diffusion and drift coefficients satisfy the "Einstein relation" which relates the values of diffusion and mobility coefficients as: $\frac{D_p}{\mu_p} = \frac{D_n}{\mu_n} = \frac{kT}{q}$. We

see cropping up again and again in semiconductor physics the "thermal voltage" kT/q which is approximately 26 millivolts at room temperature.

$x = 0$ has dropped, and the particles have spread to the left and the right. Further smearing of the particle concentration occurs at t_2 and t_3 . As time reaches infinity, the concentration is the same everywhere, and diffusion ends. A closed-form solution for this diffusion problem exists;²¹ the electron concentration everywhere is:

$$n(x,t) = \frac{A}{\sqrt{4\pi D_n t}} e^{-\frac{x^2}{4D_n t}} + n_o \quad [3-14]$$

where A is a constant, D_n is the diffusion constant for electrons, and n_o is the equilibrium concentration of electrons.

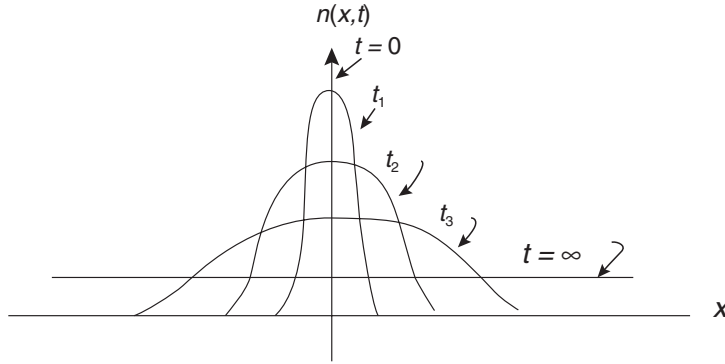


Figure 3-7: Illustration of a diffusion thought experiment.
At time $t = 0$ there is a very high concentration of particles at $x = 0$.
The particle concentration varies with time as shown, with $t_3 > t_2 > t_1$, etc.

In a famous experiment devised in 1949 called the Shockley-Haynes experiment²² (**Figure 3-8a**) an area in a piece of semiconductor material was illuminated while an electric field E_x was applied. The electric field is proportional to the voltage applied to the left side of the sample, and the electric field points from left to right.

At time $t = 0$, a light source is turned on and creates a high concentration P_o of holes at $x = 0$. The light is turned off and the concentration of holes diffuses due to the concentration gradient and drifts to the right due to the applied electric field. A detector is used to measure the hole concentration at $x = L$. The solution for the hole concentration is given by:

$$p(x,t) = \frac{P_o}{2\sqrt{\pi D_p t}} e^{-\left(\frac{(x - \mu_p E_x t)^2}{4D_p t} + \frac{t}{\tau_p}\right)} + p_{no} \quad [3-15]$$

²¹ See, e.g., R. B. Adler, et. al, *Introduction to Semiconductor Physics*, SEEC Volume 1, p. 174, 1964.

²² This was the first direct observation of minority carrier drift and diffusion dynamics in semiconductors. Described in great detail in Shockley's Nobel lecture (1956) and in J. R. Haynes and W. Shockley, *Physical Review*, vol. 75, p. 691 (1949), and J. R. Haynes and W. Shockley, "The Mobility and Life of Injected Holes and Electrons in Germanium," *Physical Review*, vol. 81, p. 835, 1951. It's also described in SEEC Volume 1, referenced earlier.

Chapter 3

The evolution of the concentration of holes in the sample is shown in **Figure 3-8b**. Using this result we can find the drift coefficient for holes μ_p by viewing how fast the pulse center moves from left to right. The drift coefficient is:

$$\mu_p = \frac{v_x}{E_x} \quad [3-16]$$

where v_x is the velocity of the center of the pulse from left to right. The diffusion coefficient D_p can be found by viewing how fast the pulse spreads.

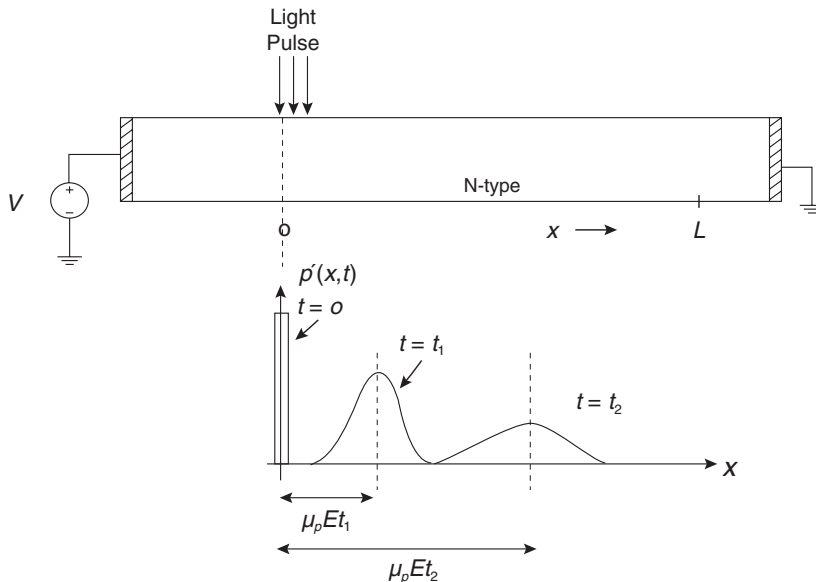


Figure 3-8: Illustration of the Haynes-Shockley experiment. The concentration of extra holes ($p'(x,t)$) moves to the right with speed $\mu_p E$ where μ_p is the mobility of the holes.

Generation and recombination

What happens when a free electron traveling through the lattice happens to meet a hole? The electron will tend to fill the hole, and the electron and hole effectively annihilate each other. This process is known as *recombination*. Similarly, the process by which carrier pairs are created is called *generation*. Carrier pairs can be created by heating up the lattice or by illuminating the semiconductor with light. It must be understood that in equilibrium in a semiconductor, the processes of generation and recombination are always going on and are in balance.

Comment on total current in semiconductors

In a semiconductor, the total current flowing is a combination of the currents due to drift plus the currents due to diffusion. What makes this a complicated process is that in a given transistor, there are drift and diffusion components due to both holes and electrons. So, there are four components of total current that we need to consider. We'll look at how this works later on, but for now let's consider the effects of semiconductor doping.

Effects of Semiconductor Doping

When a semiconductor material is doped, very small quantities of selected impurities are selectively added which significantly alters the balance of holes and electrons that exists in an intrinsic semiconductor.

Donor doped material

Typical dopants for N-type semiconductors are periodic table (**Figure 3-9**) column V elements such as phosphorus and arsenic. These column V elements have five outer-shell electrons. When, for instance, a phosphorus atom displaces a silicon atom in the lattice, there is one extra outer shell electron available; this extra electron is fairly free to move about the lattice, and acts as a mobile particle with a net negative charge of $-q$.

Figure 3-9: Section of the periodic table, showing column III and column V elements which are commonly used as dopants in semiconductors. "Acceptor" dopants are from column III and "donor" dopants are from column V.

	III	IV	V	VI
	5 B	6 C	7 N	8 O
	13 Al	14 Si	15 P	16 S
IIB				
30 Zn	31 Ga	32 Ge	33 As	34 Se
48 Cd	49 In	50 Sn	51 Sb	52 Te

Doping silicon with phosphorus or arsenic increases the free-electron population above the intrinsic electron concentration n_i . Another effect is to lower the equilibrium concentration of holes; for high doping levels, there are lots of free electrons available to recombine with holes and hence the population of holes goes down. The net effect of all of this is that in an N-type material at equilibrium with significant doping²³ ($N_D \gg n_i$) the equilibrium carrier concentrations are:²⁴

$$\begin{aligned} n_{no} &\approx N_D \\ p_{no} &\approx \frac{n_i^2}{N_D} \end{aligned} \quad [3-17]$$

²³ This assumes that the doping is not so high that the semiconductor becomes *degenerate*. For a discussion of degenerate semiconductor statistics, see, e.g., Robert F. Pierret, *Modular Series on Solid State Devices*, volume 1, "Semiconductor Fundamentals," published by Addison-Wesley, Reading Mass., 1983.

²⁴ In this terminology, n_{no} is the equilibrium concentration of free electrons in an N-type material, and p_{no} is the equilibrium concentration of holes in N-type material. This same nomenclature is used in Paul Gray and Campbell Searle, *Electronic Principles Physics, Models and Circuits*, published by John Wiley, 1967.

Acceptor doped material

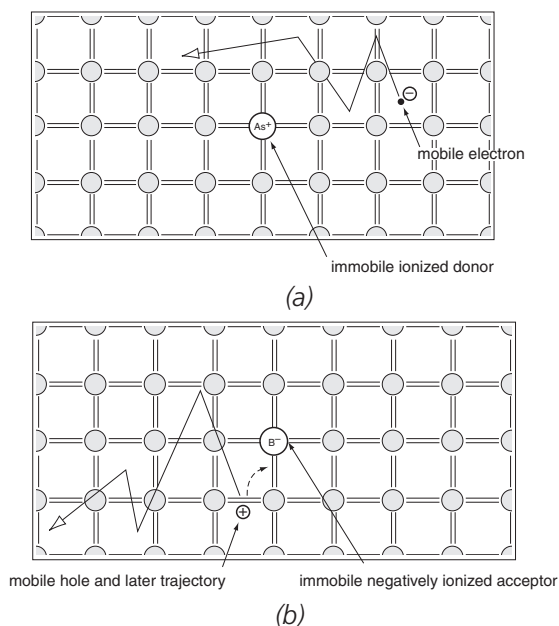
Boron is in column III of the periodic table of elements and hence has three outer-shell electrons and is a typical dopant for P-type semiconductors. Boron is therefore an “acceptor” donor. When a boron atom displaces a silicon atom in the lattice, there is one fewer outer shell electron available to generate covalent bonds. The result is that there is a quantum “void” called a “hole” which acts as a mobile particle with a net positive charge. For P-type semiconductors the equilibrium hole and electron concentrations are:

$$\begin{aligned} p_{po} &\approx N_A \\ n_{po} &\approx \frac{n_i^2}{N_A} \end{aligned} \quad [3-18]$$

Figure 3-10: Donor and acceptor doped semiconductors.

(a) Donor doped semiconductor, showing free electrons which can contribute to current flow.

(b) Acceptor doped semiconductor, with free hole.



PN Junction Under Thermal Equilibrium

We can now put together these concepts of carrier flow, creation and annihilation—that is, drift, diffusion, generation and recombination—to show how a diode operates. First we’ll consider the case of “thermal equilibrium”; the diode has no connections, no current flow, and is sitting on a lab bench. Furthermore, there is no external excitation on the diode, such as a light source.

Let’s consider an idealized PN diode structure. A one-dimensional model of a bipolar junction diode, comprised of adjacent P and N regions with conducting contacts on either end is shown in **Figure 3-11a**, shown also with the diode circuit model. The anode of the diode is the connection to the P-region, and the cathode is the connection to the N-region. We’ll discuss the “depletion region” as shown in the diagram near the PN junction later on. Let’s

consider an “abrupt” junction²⁵ where the doping changes abruptly at $x = 0$ as shown in **Figure 3-11b**. On the P-side of the junction is a concentration of N_A acceptors, and on the N-side of the junction is a concentration of N_D donors.



Figure 3-11: Ideal PN junction geometry. (a) Diode, showing anode side (P-side) and cathode side (N-side). (b) Doping levels on either side of the junction, with N_A acceptors on the P-side and N_D donors on the N-side. Near the junction is a “depletion region.”

Initially, let’s consider the behavior of this diode without any contacts at either end, and hence no applied voltage. Let’s do a thought experiment and consider what happens if we bring P-material in close proximity to the N-material. When the P-material is brought in proximity with the N-material, the high concentration of holes in the P material tends to diffuse into the N-material; the high concentration of electrons in the N-material tends to diffuse into the P-material.

If diffusion was the only carrier motion process, eventually the concentrations of electrons and holes would be the same throughout the entire diode. However, when holes diffuse from the P-side to the N-side, fixed negative charges in the silicon lattice are uncovered. When electrons diffuse from the N-side to the P-side, fixed positive charges in the lattice are uncovered. Hence, near the junction there is a net negative charge on the P-side. By the same reasoning, there is a net positive charge near the junction on the N-side. The fixed negative charge on the P-side tends to oppose further diffusion of electrons from the N-side to the P-side. Likewise, the fixed positive charges on the N-side tend to oppose diffusion of holes from the P-side to the N-side. As we’ll see later on, we can quantify this effect by noting that a barrier electric field is set up which opposes diffusion. The equilibrium concentration of electrons and holes throughout the diode is shown graphically in **Figure 3-12**.

²⁵ This is an approximation to a junction where the doping levels change very abruptly from P-type to N-type at $x = 0$.

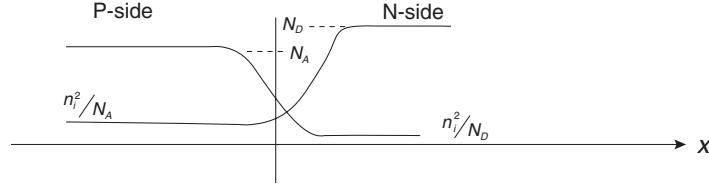


Figure 3-12: Carrier concentrations in a PN junction diode without any applied voltage bias. On the N-side, the concentration of electrons is approximately N_D far away from the barrier, and the concentration of holes is n_i^2/N_D . On the P-side, the concentration of holes far away from the barrier is N_A and the concentration of electrons far from the barrier is n_i^2/N_A .

This net charge near the junction creates an electric field. This field acts as a barrier to further current flow across the junction. The charge profile, electric field and potential for this junction at equilibrium are shown in **Figure 3-13**.²⁶ In equilibrium, there is a delicate balance between drift and diffusion in the diode. There are four current components (drift and diffusion for electrons and holes), but in the end, there is no diode current.

This drift-diffusion process results in a region near the junction that is nearly devoid of free carriers. This so-called depletion region extends from $-x_p$ to $+x_n$ as shown in the diagrams.

We can find the electric field $E(x)$ by using the one-dimensional form of Gauss's law, which is:

$$\frac{dE(x)}{dx} = \frac{-\rho(x)}{\epsilon_{Si}} \quad [3-19]$$

where $\rho(x)$ is the net charge and ϵ_{Si} is the dielectric permittivity of the silicon material. Since the charge $\rho(x)$ is piecewise constant on the P and N sides, the electric field is piecewise linear on the P and N sides, as shown. The maximum electric field is found at the junction at $x = 0$ and is given by:

$$|E_{\max}| = \frac{qN_A x_p}{\epsilon_{Si}} = \frac{qN_D x_n}{\epsilon_{Si}} \quad [3-20]$$

We can find the junction potential similarly by integrating the electric field, since

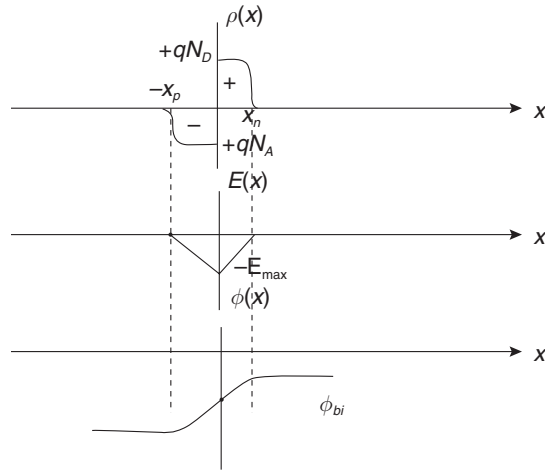
$$\frac{d\phi(x)}{dx} = -E(x) \quad [3-21]$$

At zero bias, there is a voltage ϕ_{bi} , called the *built-in potential*, across the depletion region of the junction. This voltage is a result of an electric field that opposes the diffusion of mobile electrons and holes across the junction. The built-in potential is given by:

$$\phi_{bi} = \frac{kT}{q} \ln \left(\frac{N_A N_D}{n_i^2} \right) \quad [3-22]$$

²⁶ Note that these pictures are drawn for an abrupt junction, where it is assumed that doping on the N-side is N_D and doping on the P-side is N_A .

Figure 3-13: Ideal PN junction in thermal equilibrium, in a one-dimensional approximation. These plots show the depletion region charge density $\rho(x)$, electric field $E(x)$ and junction potential $\phi(x)$.



We'll see that the term kT/q comes up over and over in semiconductors, where k is Boltzmann's constant, T is absolute temperature in Kelvin, and q is the electronic charge. This "thermal voltage" is approximately 26 millivolts at room temperature.

PN Junction Under Applied Forward Bias

When we forward bias a diode, we apply a positive voltage between the anode (P-side contact) and the cathode (N-side contact) of the diode. Under forward bias, the delicate balance between drift and diffusion is altered significantly and net current flows in the diode. The application of a forward-bias voltage (i.e., a positive potential applied to the P-side relative to the N-side) reduces the amplitude of the barrier electric field at the junction. When this barrier is reduced, it is easier to inject holes from the P-side to the N-side and electrons from the N-side to the P-side. This causes a net forward current to flow.

Shown in **Figure 3-14** are the junction charge, electric field and voltage under forward ($V > 0$), reverse ($V < 0$) and no-bias ($V = 0$) conditions. When there is no bias—i.e., no voltage applied to the diode—the voltage across the diode depletion region is just the built-in potential ϕ_{bi} . There is no net diode current because all components of drift and diffusion current in the diode exactly cancel each other. The electric field barrier at the junction is as shown.

Under reverse-bias, the applied voltage adds to the junction built-in potential ϕ_{bi} . This in turn means that, under reverse bias, the electric field barrier is higher, and the depletion region is wider. Under higher reverse bias, the depletion region width increases and the maximum electric field increases as well. If you keep increasing the reverse bias, the electric field increases until junction breakdown is reached. Junction breakdown in silicon occurs at an electric field on the order of 3×10^5 V/cm.

Under forward bias, the electric field barrier is decreased by the applied voltage. This reduction in electric field barrier allows holes to be injected from the P-side to the N-side, and electrons to be injected from the N-side to the P-side. The injected carriers then diffuse away from the junction, setting up a net diode current—called the diode *forward* current. As we'll see later on, there's an exponential relationship between diode voltage and current when the diode is forward-biased.

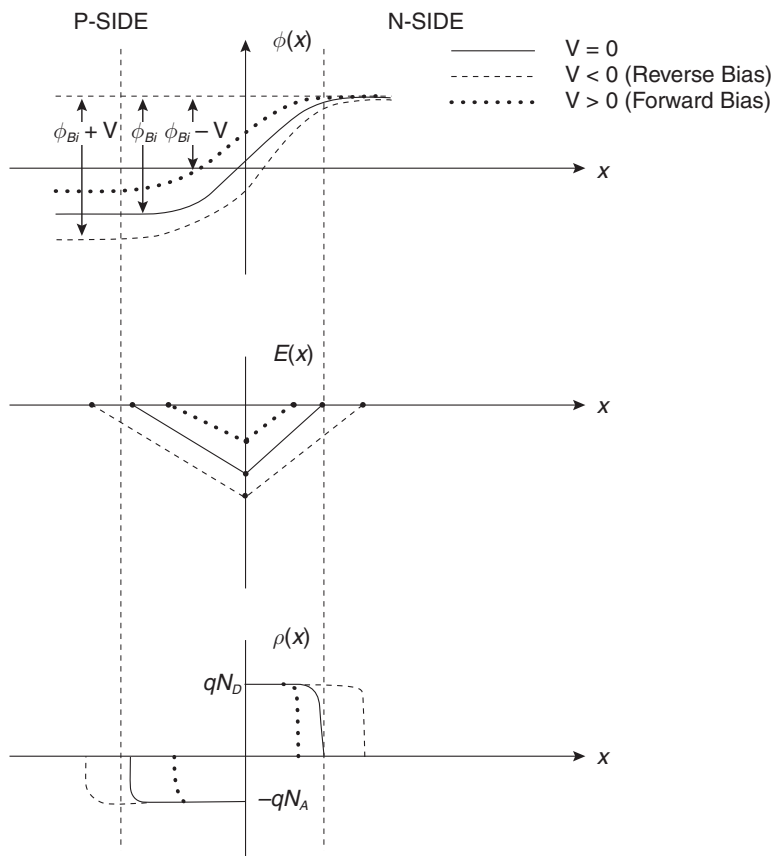


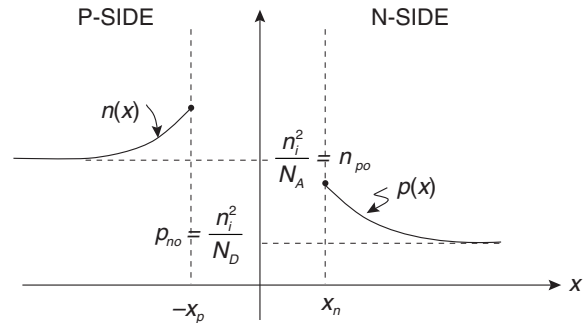
Figure 3-14: Depletion region under zero bias, forward and reverse bias showing depletion region charge density $\rho(x)$, electric field $E(x)$ and junction potential $\phi(x)$. Under reverse bias the width of the depletion region grows and the maximum electric field grows as well. Under forward bias, the converse is true.

Shown in **Figure 3-15** are the concentrations of minority carriers²⁷ on the N-side and on the P-side of the semiconductor, with the diode under forward bias. Under forward bias, the electric-field barrier due to the depletion region is lowered, and electrons are injected from the N-side to the P-side, and holes are injected from the P-side to the N-side. The injected electrons on the P-side diffuse to the left. The injected holes on the N-side diffuse to the right. This diffusion results in a net diode current from left to right. Detailed considerations²⁸ show us that there is an exponential relationship between diode voltage and the excess minority carrier concentrations at the edge of the depletion regions, or:

$$\begin{aligned} n'(-x_p) &= n(x) - n_{po} = A \left(e^{\frac{qV_D}{kT}} - 1 \right) \\ p'(x_n) &= p(x) - p_{no} = B \left(e^{\frac{qV_D}{kT}} - 1 \right) \end{aligned} \quad [3-23]$$

where A and B are constants having to do with doping levels and the intrinsic carrier concentration,²⁹ n_{po} is the electron concentration on the P-side far from the junction, and p_{no} is the hole concentration on the N-side far from the junction.

Figure 3-15: Minority carrier concentrations in a PN junction under forward bias. n_{po} is the concentration of electrons far from the junction on the P-side. p_{no} is the concentration of holes far from the junction on the N-side.



²⁷ Let's discuss why the carriers shown are "excess minority carriers." Consider the concentration of n' electrons on the P-side of the diode. The n' electrons are minority carriers since electrons are in the minority on the P-side of the diode. This, of course, assumes that the number of carriers injected is small compared to the majority carrier concentration N_A , or $n' \ll N_A$. This condition is termed in the literature *low-level injection*. We call the plotted carriers n' "excess" because these carriers are in excess of the equilibrium concentration n on the P-side. Remember that the equilibrium concentrations of carriers on the P-side are $p \approx N_A$ and $n \approx n_i^2/N_A$.

²⁸ See, e.g., Pierret's *Modular Series on Semiconductor Devices, Volume II, The PN Junction Diode*, pp. 48–50. The factor kT/q (with k = Boltzmann's constant, T = absolute temperature in Kelvin and q = electronic charge) comes up again and again in semiconductors. The value of kT/q is approximately 26 millivolts at room temperature (300K).

²⁹ Again, detailed considerations show that $A = n_i^2/N_A$ and $B = n_i^2/N_D$, where N_A is the acceptor concentration on the P-side and N_D is the donor concentration on the N-side.

We're now in a position to piece together arguments for the exponential voltage-current relationship in a forward-biased diode. The keys to realizing this are as follows:

- Holes are injected from the P-side to the N-side, and electrons are injected from the N-side to the P-side.
- These injected holes result in a concentration gradient on either side of the depletion region.
- The excess minority carrier concentration at the edge of the depletion region follows an exponential relationship with the diode voltage V_D .
- This concentration gradient results in a diffusion current of excess minority carriers. In essence, the injected carriers diffuse away from the junction. As the carriers diffuse away, they recombine with majority carriers. The shape of the curve follows an exponential decay away from the junction as:³⁰

$$p'(x) = p(x = x_n) e^{\frac{-x'}{L_p}}$$

- The minority carrier currents are dominated by diffusion. In other words, the minority diffusion current is larger than the minority drift current. This is due to the fact that the minority concentration is small (shown before due to the law of mass action). However, the minority concentration gradient is large and diffusion dominates.

The individual current components in a forward-biased diode are shown in **Figure 3-16**. Let's first focus on the electron current on the P-side, labeled J_e . We know that there's a concentration gradient of electrons near the depletion region, since these charges are injected from the P-side. This concentration gradient causes a diffusion current to flow. In essence, electrons diffuse away from the junction (to the left in the diagram). This is the current J_e as shown. As the excess carriers diffuse away, they recombine with holes in the p-bulk material. This causes a net hole current J_h as shown.

It is assumed that the electron current J_e and the hole current J_h are constant through the depletion region. We assume that the depletion region is so depleted of charges that there's no significant recombination in this region.

Now, what about the exponential diode voltage-current relationship? Since we know that the total diode current is the sum of electron and hole currents, and since we assume that electron and hole current is constant through the depletion region, we can write that the diode current is:

$$J_{total} = J_e(x = -x_p) + J_p(x = x_n) \quad [3-24]$$

³⁰ L_p is the minority carrier diffusion length for holes. It's a measure of how far the holes diffuse into the n-bulk material before recombining. Numerically, $L_p = \sqrt{D_h \tau_h}$ where D_h is the diffusion constant for holes and τ_h is the minority carrier lifetime, or the characteristic lifetime for decay of excess carriers.

The individual components of current $J_e(x = -x_p)$ and $J_p(x = x_n)$ are dominated by diffusion, and we know that the concentration gradient depends on the junction voltage, or:

$$\begin{aligned} J_e(x = -x_p) &\propto e^{\frac{qV_D}{kT}} \\ J_p(x = x_n) &\propto e^{\frac{qV_D}{kT}} \end{aligned} \quad [3-25]$$

Therefore, we see the derivation of the exponential voltage-current relationship for the ideal diode. The electron current (J_e), the hole current (J_h) and the total current through the diode is shown in **Figure 3-16**.

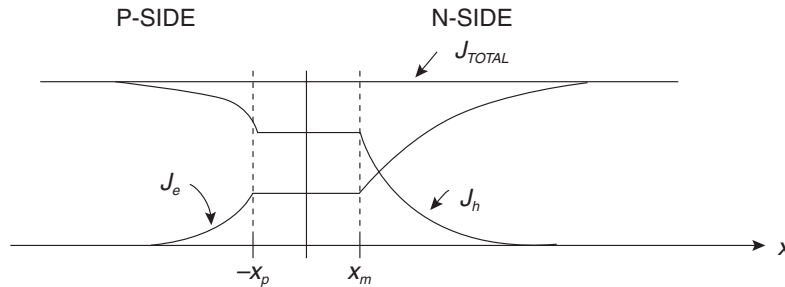


Figure 3-16: Various current components in a diode under forward-bias.

J_e is the current density (A/m^2) due to electron flow, J_h is the current density due to hole flow, and J_{total} is total current density.

Reverse Biased Diode

Under reverse bias, the minority carrier concentration at the edges of the depletion region are suppressed below their equilibrium values, as shown in **Figure 3-17**. As shown before, there is also a large electric field barrier that suppresses current flow across the junction. As we'll see in the next section, the diode current under reverse bias is small, but finite.

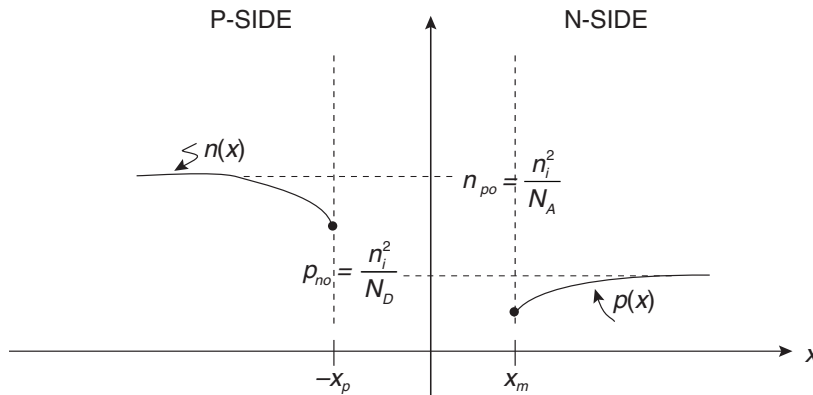


Figure 3-17: Minority carrier concentrations under reverse bias.

Ideal Diode Equation

We showed before that the current densities J_e and J_h , measured at the edge of the depletion region, each have a value that depends exponentially on diode voltage. Since the total current is the sum of these two components, the total diode current has this same functional dependence. This result is summarized by the familiar Shockley equation, which is:

$$I_D = J_{TOTAL} A = I_s \left(e^{\frac{qV_D}{kT}} - 1 \right) \quad [3-26]$$

where J_{TOTAL} is the current density (in amps/m²), A is junction area, and I_s is the *saturation current* with typical values of 10⁻¹⁵ amperes for a signal diode. Under forward bias, with $V_D \gg kT/q$, the diode current is approximately:

$$I_D \approx I_s e^{\frac{qV_D}{kT}} \text{ for } V_D \gg kT/q \quad [3-27]$$

For reverse voltages $V_D \ll -kT/q$ the diode current saturates at the reverse saturation current $-I_s$.

The total plot of diode current as a function of voltage (I_D vs. V_D) for the ideal diode is shown in **Figure 3-18a**. An often-used approximate curve is shown in **Figure 3-18b**. Sometimes the more realistic idealized curve of **Figure 3-18c** is used, which incorporates the diode “knee” voltage of approximately 0.6 volts. You can also include the effects of any diode ohmic resistance (not accounted for by ideal diode analysis) by including a finite slope in the diode curve, as shown in **Figure 3-18d**.

Detailed considerations show that the ideal diode V_D/I_D relationship is given by:

$$I_D = qAn_i^2 \left(\frac{D_p}{N_D L_p} + \frac{D_n}{N_A L_n} \right) \left(e^{\frac{qV_D}{kT}} - 1 \right) = I_s \left(e^{\frac{qV_D}{kT}} - 1 \right) \quad [3-28]$$

where A is the junction area, D_p and D_n are diffusion constants for holes and electrons, and L_p and L_n are diffusion lengths for holes and electrons. The term I_s is the reverse saturation current of the diode.

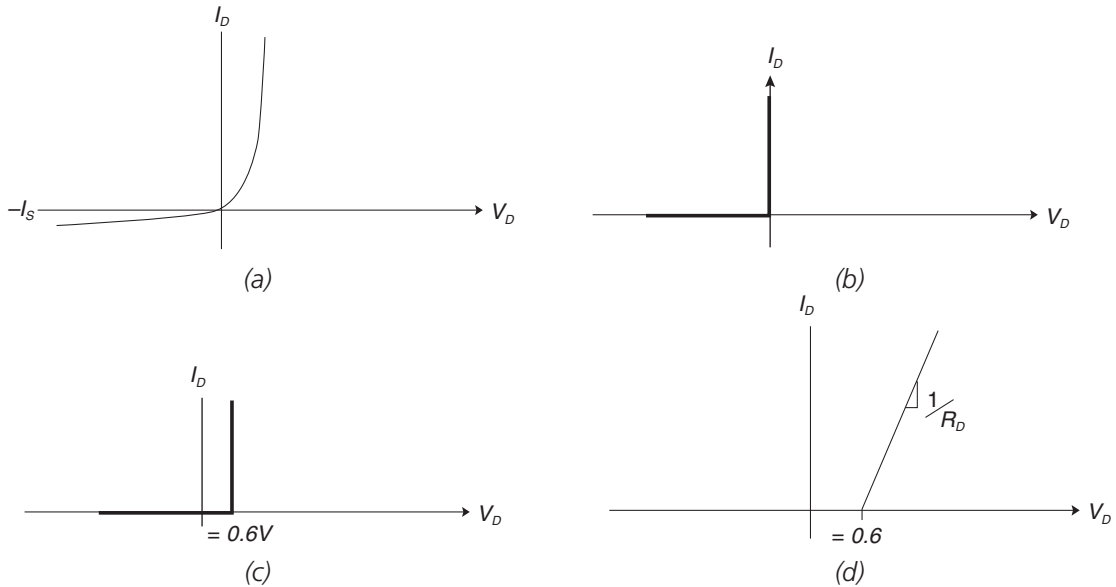


Figure 3-18. Plot of I_D vs. V_D for diode, showing various approximations. (a) Result for ideal diode showing exponential V-I curve. (b) Often-used simple approximation where the diode turns on with zero voltage drop. (c) More realistic approximation including diode ON voltage. (d) Including diode ohmic resistance R_D .

Charge Storage in Diodes

Charge storage in diodes is accomplished via several mechanisms. First, there is charge storage in the depletion regions, under both forward and reverse bias conditions. This charge is stored in the dipole layer near the junction.

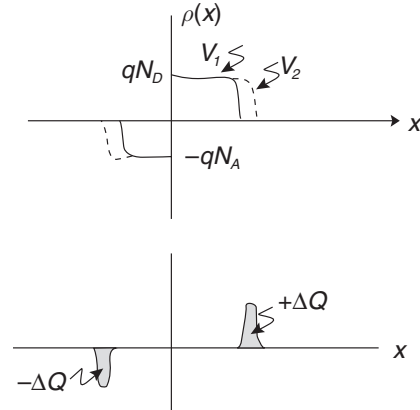
Under forward bias, charge is stored in the neutral regions of the P-side and the N-side. We've seen that carriers are injected from one side of the junction to the other and then diffuse away; therefore, there is a "diffusion capacitance" which controls how much charge is stored in the neutral regions. We'll discuss each of these charge storage mechanisms next.

Depletion capacitance

We've seen earlier that a variation in applied diode bias results in a change in the depletion width, and hence a change in the charge in the depletion region. Let's revisit this from a graphical point of view, as shown in **Figure 3-19**. When the reverse bias on the diode increases from V_1 to V_2 , the depletion region width increases. We can think of this as a change in charge stored in the depletion region. There is a net change in charge of $\Delta Q+$ on the N-side, and a change in charge of $\Delta Q-$ on the P-side. Capacitance is equal to the derivative of charge with respect to voltage, so we can find the junction capacitance as:

$$C_j = \frac{dQ}{dV} \quad [3-29]$$

Figure 3-19: Illustration of cause of diode junction capacitance. As the reverse bias increases from V_1 to V_2 , the depletion width increases resulting in an increase in depletion charge $+\Delta Q$ and $-\Delta Q$ on either sides of the junction.



A detailed analysis shows that the junction capacitance can be expressed as:

$$C_j = \frac{C_{jo}}{\left(1 - \frac{V_D}{\phi_{bi}}\right)^m} \quad [3-30]$$

where $m = 1/2$ for an abrupt junction and $m = 1/3$ for a linearly graded junction. The term C_{jo} is the diode capacitance at zero diode voltage, and ϕ_{bi} is the built-in voltage. The junction voltage for a typical signal diode is plotted in **Figure 3-20**. Note that as reverse-bias voltage increases, the junction capacitance decreases.³¹

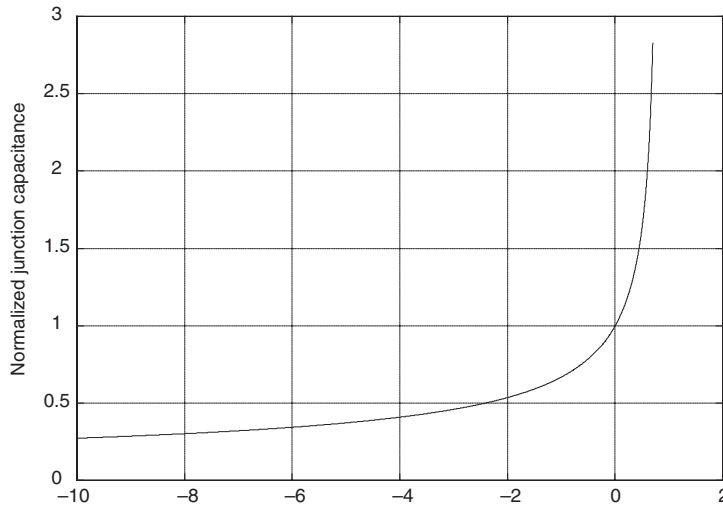


Figure 3-20: Nonlinear junction capacitance vs. diode voltage for abrupt junction diode, shown for $\phi_{bi} = 0.8V$ and $C_{jo} = 1$.

³¹ This simple junction capacitance equation works well assuming that $V_D < \phi_{bi}$. Note that this equation shows a junction capacitance that increases without bound as $V_D \rightarrow \phi_{bi}$.

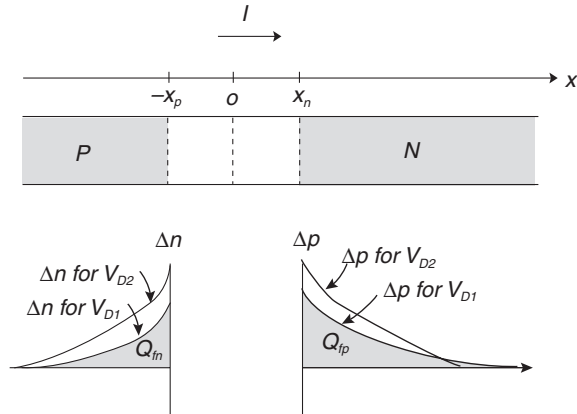
Charge Storage in the Diode Under Forward Bias

When the diode is forward biased and carrying forward current, there is charge stored in the neutral P and N regions. If the diode current increases, this charge must likewise increase. As the voltage applied to the diode increases, the diode current increases, and the charge stored in the *diffusion capacitance* increases as well. These effects are illustrated in **Figure 3-21**.

The charge stored in the P-region is the area under the Δn curve and is called Q_{fn} . Likewise, the charge stored in the N-region is the area under the Δp curve, and is called Q_{fp} . Let's say that the diode forward voltage is increased from V_{D1} to V_{D2} . The height curves for the stored charges Δn and Δp also increase because more charges are injected through the depletion region. Correspondingly, the total stored charge in the diode (the sum of Q_{fn} and Q_{fp}) increases as the diode voltage increases from V_{D1} to V_{D2} . A change in charge with a change in diode voltage can be modeled as an equivalent capacitance. This capacitance is sometimes called the *diffusion capacitance* of the forward biased diode.

These charges diffuse away from the junction and contribute to total diode current I as shown. Note that electrons diffuse from right to left, and holes diffuse from left to right, yielding a net current from left to right.

Figure 3-21: Excess minority carrier concentrations under forward bias, short diode. As the forward diode voltage increases from V_{D1} to V_{D2} , the electron and hole concentrations increase as shown.



The result of this is that the diffusion capacitance is proportional to the diode forward current, or:

$$C_d \propto I_D \quad [3-31]$$

The details of the proportionality constant depend on the specific construction of the diode.

Reverse Recovery in Bipolar Diodes

When a bipolar diode is ON and you abruptly reverse the voltage across the diode, you might expect the diode current to immediately drop to zero. However, if you perform this experiment and monitor the diode current, you'll find that the diode current actually spikes negative for a period of time, called the *reverse recovery* time. This reverse recovery is a direct result of the charge stored in the diffusion capacitance of the diode. This charge must be removed before the diode can turn off.

An experiment illustrating reverse recovery is shown in **Figure 3-22**. A diode is repetitively pulsed by a voltage source $v(t)$ through a resistor R . Let's assume that the amplitude V_o of the pulse is much larger than the diode ON voltage of $\sim 0.7\text{V}$. When the input voltage transitions positive, the diode current reaches a limit of $\sim V_o/R$, limited by the resistance.

Reverse recovery occurs after the input voltage transitions from $+V_o$ to $-V_o$ at time $t = T/2$. Looking at the diode current curve $i_d(t)$ we see that the diode current is negative for a time period t_{rr} , called the *reverse recovery time*. During this reverse recovery time, the charge stored in the diffusion capacitance is removed, turning the diode off.

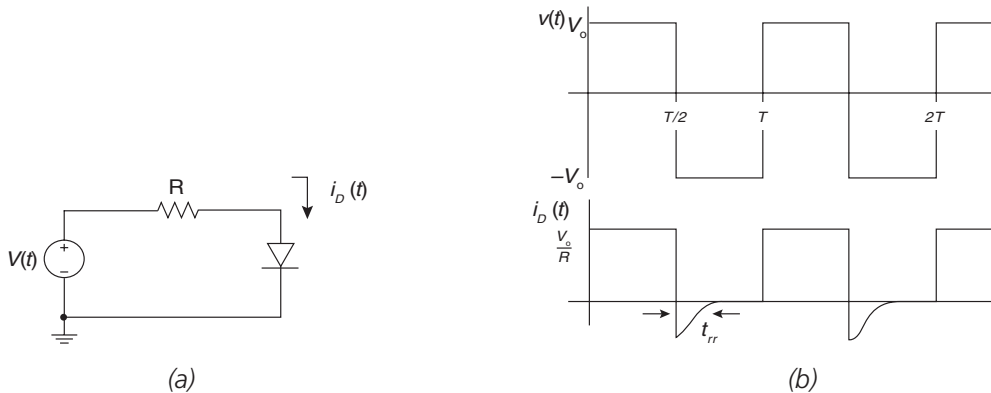


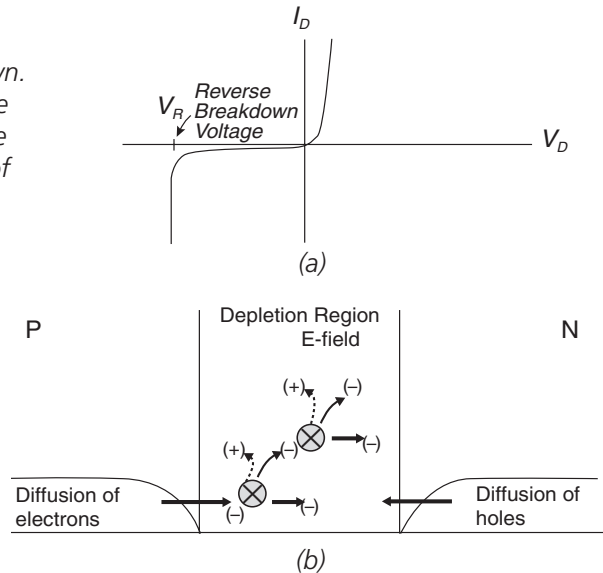
Figure 3-22: Test circuit illustrating reverse recovery. (a) Test circuit. (b) Waveforms showing driving voltage $v(t)$ and resultant diode current $i_D(t)$. When the driving voltage transitions negative, the diode current goes negative for a time t_{rr} (the reverse recovery time). During this reverse recovery interval, extra charges are swept out of the diode neutral regions as the diode shuts off.

Reverse Breakdown

A reverse-biased diode carries a small reverse current. This is true until a large-enough reverse voltage, called the reverse-breakdown voltage, is applied (**Figure 3-23a**). When the reverse breakdown voltage is applied, the reverse current carried by the junction increases significantly. If there is not a current-limiting mechanism, this reverse current can destroy the device.

In signal diodes, reverse breakdown is dominated by an effect known as *avalanche breakdown*. In a reverse-biased junction, the applied reverse voltage applies an electric field in the depletion region of the diode. At a critical field E_{max} the carriers in the depletion region gather sufficient speed to hit the lattice with enough force to cause other electron-hole pairs to be created. These created electron-hole pairs are then accelerated in the electric field, creating other electron-hole pairs. This process results in a rapidly increasing diode current as the reverse voltage is increased.

Figure 3-23: Diode reverse breakdown. (a) Diode V-I curve, showing reverse breakdown when the diode reverse voltage reaches V_R . (b) Illustration of reverse breakdown process.



Taking a Look at a Diode Datasheet

In this section, we'll examine the datasheet of a 1N914 diode,³² an inexpensive high-speed switching diode. The maximum ratings of the diode (**Figure 3-24**) show that this diode has a reverse voltage rating of 100 volts and a forward current of 200 milliamps. The maximum peak current rating is 500 milliamps for a short period of time. These numbers are typical of a small signal switching diode.

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Reverse Voltage	V_R	100	Vdc
Forward Current	I_F	200	mAdc
Peak Forward Surge Current	$I_{FM}(\text{surge})$	500	mAdc

Figure 3-24: Maximum ratings of 1N914 diode.

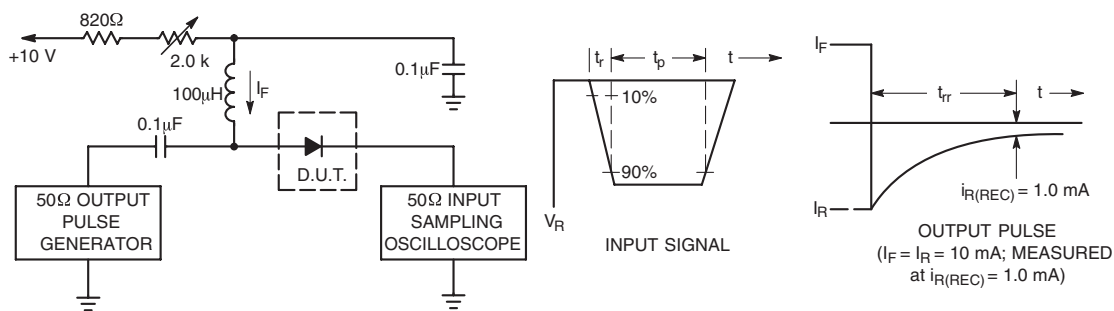
Next, let's look at the basic electrical characteristics (**Figure 3-25**). Again, we see the reverse voltage rating ("Reverse Breakdown Voltage") of 100V. The reverse leakage current is specified to be a maximum of 5 mA at a reverse voltage of 75V. We also see information on the diode capacitance; the capacitance information is better displayed later on in a graph. We also see that this diode has a fast reverse recovery time of $t_{rr} < 4$ ns, when tested with the test circuit shown in **Figure 3-26**.

³² The particular datasheet excerpts are from the On Semiconductor MMDL914 diode, the surface mount version of the through-hole 1N914 diode. Reprinted with permission of On Semiconductor.

Figure 3-25: Electrical characteristics of 1N914 diode.

ELECTRICAL CHARACTERISTICS ($T_A = 25^\circ\text{C}$ unless otherwise noted)

Characteristic	Symbol	Min	Max	Unit
OFF CHARACTERISTICS				
Reverse Breakdown Voltage ($I_R = 100\ \mu\text{A}$ dc)	$V_{(BR)}$	100	–	Vdc
Reverse Voltage Leakage Current ($V_R = 20\ \text{Vdc}$) ($V_R = 75\ \text{Vdc}$)	I_R	– –	25 5.0	nAdc μA dc
Diode Capacitance ($V_R = 0\ \text{V}$, $f = 1.0\ \text{MHz}$)	C_T	–	4.0	pF
Forward Voltage ($I_F = 10\ \text{mA}$ dc)	V_F	–	1.0	Vdc
Reverse Recovery Time ($I_F = I_R = 10\ \text{mA}$ dc) (Figure 1)	t_{rr}	–	4.0	ns



- Notes: 1. A 2.0 kΩ variable resistor adjusted for a Forward Current (I_F) of 10 mA.
2. Input pulse is adjusted so $I_{R(\text{peak})}$ is equal to 10 mA.
3. $t_p \gg t_{rr}$

Figure 3-26: Reverse recovery time test circuit for 1N914 diode.

The forward and reverse characteristics of the diode at various temperatures are shown in **Figure 3-27**. The forward voltage characteristic shows a negative temperature coefficient of diode voltage of approximately $-2\ \text{mV}$ per degree Celsius. At constant collector current, the diode voltage decreases approximately this much as temperature increases. The forward plots also show additional curvature at high diode currents, due to ohmic drops in the diode not accounted for by the ideal diode model.

Shown in the reverse leakage plot is the fact that reverse leakage current significantly increases as temperature increases.

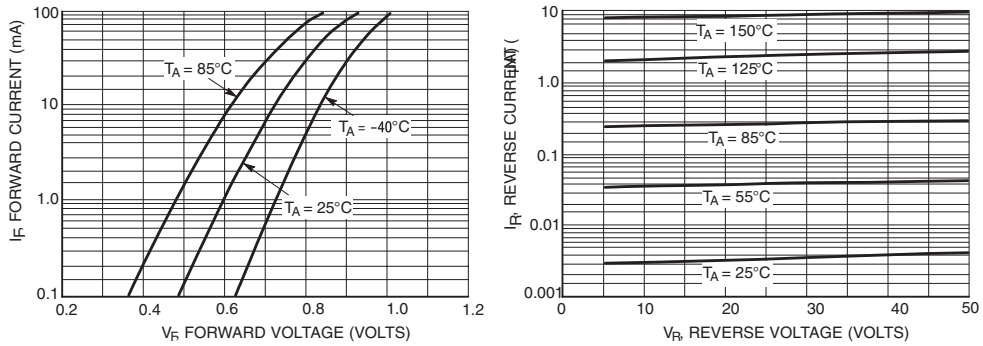


Figure 3-27: Forward and reverse characteristics of 1N914 diode.

The chart of diode capacitance in the reverse region (**Figure 3-28**) shows that the diode capacitance is nonlinear and decreases as reverse bias voltage increases.

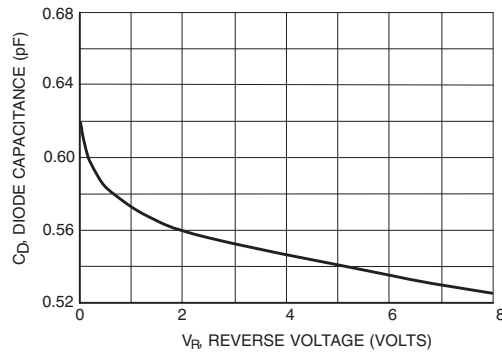


Figure 3-28: Diode capacitance in reverse region for 1N914 diode.

The thermal characteristics (**Figure 3-29**) show that the maximum device dissipation is 200 milliwatts at an ambient temperature of 25°C . We need to derate this maximum power dissipation if the ambient temperature is higher than 25°C .

Thermal Characteristics

Characteristic	Symbol	Max	Unit
Total Device Dissipation FR-5 Board $T_A = 25^\circ\text{C}$ (Note 1) Derate above 25°C	P_D	200	mW
		1.57	mW/ $^\circ\text{C}$
Thermal Resistance, Junction-to-Ambient	$R_{\theta JA}$	635	$^\circ\text{C}/\text{W}$
Junction and Storage Temperature	T_J, T_{stg}	-55 to 150	$^\circ\text{C}$

Figure 3-29: Thermal characteristics of 1N914 diode.

Some Quick Comments on Schottky Diodes

The analysis that we have done in this chapter has focused exclusively on bipolar diodes comprised of a PN junction. Another type of diode, called a Schottky diode, is comprised of a metal semiconductor junction. We won't delve into the details but will leave the reader with a few bullets:

- The typical forward voltage of a bipolar diode at operating current is $\sim 0.6\text{V}$ or so, while the forward voltage of a Schottky is somewhat less, typically 0.4V .
- Bipolar diodes, as shown before, exhibit reverse recovery. Schottky diodes, to first order, do not have reverse recovery.

Table 3-1: Comparison of bipolar and Schottky diodes.

Item	Bipolar	Schottky
Typical forward voltage	$>0.6\text{V}$	$\sim 0.4\text{V}$
Reverse recovery?	Yes	No
Typical voltage rating	$< \text{several kV}$	$< 100\text{V}$

Chapter 3 Problems

Problem 3.1

For the ideal diode under forward bias, find the amount of forward voltage increase that causes a 10× increase in forward current at room temperature.

Problem 3.2

An abrupt PN⁺ junction diode has $N_D = 10^{16}/\text{cm}^3$ and $N_A = 10^{15}/\text{cm}^3$. Find the breakdown voltage if the breakdown electric field $E_{\max} = 3 \times 10^5 \text{ V/cm}$. Assume the abrupt junction approximation, an intrinsic carrier concentration in silicon of $n_i = 1.5 \times 10^{10}/\text{cm}^3$, that the diode operates at room temperature, and note that the total depletion region width with an applied bias V_D is:

$$x_n + x_p = \sqrt{\frac{2\varepsilon_{\text{Si}}(\phi_{bi} - V_D)}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)} \quad [3-32]$$

Problem 3.3

An abrupt PN junction has parameters $N_A = 10^{17}/\text{cm}^3$, $N_D = 10^{16}/\text{cm}^3$, and area $A = 10^{-3} \text{ cm}^2$. Find the forward voltage at a forward current $I_D = 10$ milliamps. For this diode at 300K, assume that mobilities are: $\mu_n = 1400 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$, $\mu_p = 400 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$, the minority carrier lifetimes are $\tau_n = \tau_p = 10^{-6}$ seconds, and that the intrinsic carrier concentration is $n_i = 1.4 \times 10^{10} \text{ cm}^{-3}$.

Problem 3.4

A diode has a reverse saturation current $I_S = 15$ nanoamps at room temperature (298K). Plot the diode I_D/V_D curve over the voltage range $-10\text{V} < V_D < +0.5\text{V}$, and at room temperature ($\sim 27^\circ\text{C}$, or 300K) and at elevated temperature of 150C (423K). Assume the rule-of-thumb that the reverse current of a diode doubles for every 10°C increase in temperature.

Problem 3.5

This problem investigates the use of load lines.³³ The ideal diode of Problem 3.3 above is used in a simple circuit, biased from a +12V source through a 1-kΩ resistor. Assume that the diode

³³ The load line method is a graphical method used to solve nonlinear equations. By plotting the characteristic diode curve and then plotting the curve imposed by an external resistive element and noting the intersection of the two plots we determine the operating point of the circuit. This method is also very useful for analyzing transistor circuits.

Chapter 3

operates at room temperature (300K) and hence $kT/q = 26$ mV. Find the diode current and diode voltage by the following methods:

- First, assume that the diode ON voltage is a constant 0.6V, regardless of diode current. Solve for diode current I_D given this assumption.
- In a more detailed analysis, solve iteratively the diode equation and resistor load line equation to find diode current and diode voltage, assuming that the diode reverse saturation current $I_s = 10^{-10}$ A.
- Plot the diode current/voltage load line, indicating the operating point on your plot.

Problem 3.6

Consider the diode circuit shown in **Figure 3-30**. The diode is biased with bias current source I_D , and a small signal AC current source is also connected, i_d . The diode voltage has both a DC part, V_D , and a small-signal varying part v_d . The total diode current is given by:

$$i_D = I_s \left(e^{\frac{qv_D}{kT}} - 1 \right) \quad [3-33]$$

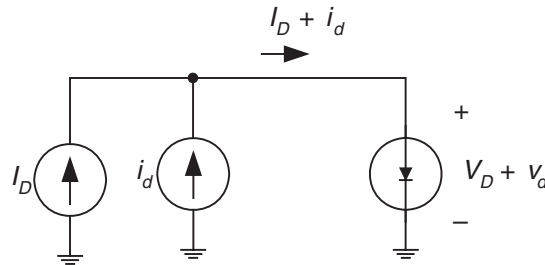
The diode current and diode voltage can be expressed as a DC part plus an incrementally varying part.

$$\begin{aligned} i_D &= I_D + i_d \\ v_D &= V_D + v_d \end{aligned} \quad [3-34]$$

Assume that the diode is biased in its forward-active region and that the value of the i_d is very small.

- Find the ratio of small-signal voltage to small-signal current, or v_d/i_d .
- Draw the low-frequency small-signal model relating v_d to i_d .

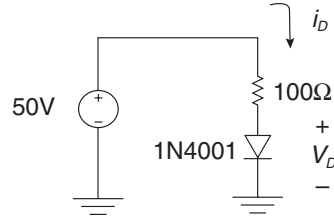
Figure 3-30: Diode circuit for Problem 3.6.



Problem 3.7

Shown in **Figure 3-31** is a circuit containing a 1N4001 diode. Draw the load line and find the diode current and diode power dissipation.

Figure 3-31: 1N4001 diode circuit for Problem 3.7.



Problem 3.8

Shown in **Figure 3-32** is a 120VAC power supply driving a circuit containing a 1N4004 diode. (A 120VAC waveform is a sinewave at 60 Hz, which has a positive peak at +170V and a negative peak at -170V). Note that in this case we use the 1N4004, since it has sufficient voltage rating for this application.

- Plot input voltage $v_{in}(t)$, diode current $i_D(t)$ and diode voltage $v_D(t)$ vs. time. You can ignore any reverse recovery.
- Calculate average power dissipated in the diode and in the resistor. In order to find power dissipation in the diode, calculate average diode current $\langle i_D \rangle$. The power dissipated in the diode³⁴ is:

$$P_{DIODE} \approx \langle V_D \rangle \langle i_D \rangle$$

where V_D is the diode ON voltage. Rather than doing a nasty integral, pick some reasonable value of $\langle V_D \rangle$ when the diode is on.

In order to find power dissipated in the resistor, we use:

$$P_{RESISTOR} \approx I_{D,RMS}^2 R$$

$$I_{D,RMS} = \sqrt{\frac{1}{T} \int_0^T (i_D(t))^2 dt}$$

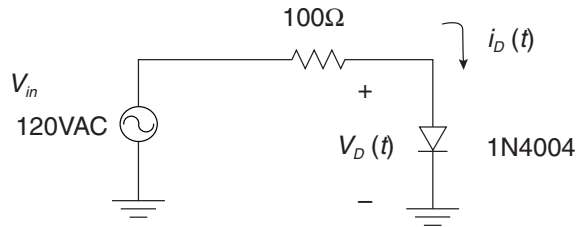


Figure 3-32: 1N4004 AC diode circuit for Problem 3.8.

³⁴ This is a very rough approximation. To get a more accurate answer we should do the integral

$$P_{DIODE} = \frac{1}{T} \int_0^T v_d(t) i_d(t) dt$$

References

- Adler, R. B., Smith, A. C., and Longini, R. L., *Introduction to Semiconductor Physics*, Semiconductor Electronics Education Committee, volume 1, John Wiley, 1964.
- Bardeen, John, “Semiconductor Research Leading to the Point Contact Transistor,” December 11, 1956, from *Nobel Lectures Physics, 1942-1962*, Elsevier Publishing, available from the Nobel e-museum at Web: <http://www.nobel.se/physics/laureates/1956/>
- Brattain, Walter H., “Surface Properties of Semiconductors,” December 11, 1956, from *Nobel Lectures Physics, 1942-1962*, Elsevier Publishing, available from the Nobel e-museum at Web: <http://www.nobel.se/physics/laureates/1956/>
- Feynman, Richard, Leighton, Ralph, and Sands, Matthew, *The Feynman Lectures on Physics*, Addison-Wesley, 1963.
- Gray, P. E., DeWitt, D., Boothroyd, A. R., and Gibbons, J. F., *Physical Electronics and Circuit Models for Transistors*, (SEEC volume 2), John Wiley, 1964.
- , *Physical Electronics and Circuit Models of Transistors*, Semiconductor Electronics Education Committee, volume 2, John Wiley, 1964. Hannay, N. B., editor, *Semiconductors*, American Chemical Society Nomograph Series, Reinhold Publishing Corp., 1959.
- Gray, Paul, and Searle, Campbell, *Electronic Principles Physics, Models and Circuits*, John Wiley, 1967.
- Haynes, J. R. and Shockley, W., “The Mobility and Life of Injected Holes and Electrons in Germanium,” *Physical Review*, vol. 81, p. 835, 1951.
- , *Physical Review*, vol. 75, p. 691, 1949.
- Pierret, Robert F., *Modular Series on Solid State Devices, Volume 1: Semiconductor Fundamentals*, Addison-Wesley, 1983.
- Shockley, William, “Transistor Electronics: Imperfections, Unipolar and Analog Transistors,” *Proceedings of the IRE*, vol. 40, no. 11, November, 1952, pp. 1289–1313, reprinted in *Proceedings of the IEEE*, vol. 85, no. 12, December 1997, pp. 2055–2080.
- , “Circuit Element Utilizing Semiconductive Materials,” U.S. Patent #2,569,347, issued September 25, 1951 and “Semiconductor Amplifier,” U.S. patent 2,502,488, issued April 5, 1950, found in “Semiconductor Amplifier Patent,” *Proceedings of the IEEE*, vol. 86, no. 1, p. 36.
- , “Transistor Technology Evokes New Physics,” December 11, 1956, from *Nobel Lectures Physics, 1942-1962*, Elsevier Publishing, available from the Nobel e-museum at Web: <http://www.nobel.se/physics/laureates/1956/>

Bipolar Transistor Models

In This Chapter

- *This chapter builds on the device model work done in the previous chapter and models for the ideal bipolar transistor are derived. Rather than deriving the full transistor equations, we will rely on results from the ideal diode and “talk through” intuitively how the bipolar transistor works. The NPN transistor is considered, but results obtained are germane to analysis of the PNP transistor as well. Most of this discussion will focus on operation of the bipolar transistor in the “forward-active” region, the region of operation where the transistor can be used as an amplifier. In later chapters, we’ll extensively discuss operation of the transistor in the saturation and cutoff regions.*

A Little Bit of History

Pioneering work on the bipolar junction transistor was done at Bell Laboratories¹ in the late 1940s, with contributions from William Shockley, William Brattain, James Bardeen, and others. This team (led by Shockley) was challenged by Bell management to invent a solid-state switch to replace mechanical relays. They initially focused their efforts on devices fabricated from germanium. The first device that demonstrated the transistor effect and current gain was the “point contact” transistor demonstrated in 1948² (**Figure 4-1**), which had a germanium crystal with closely spaced gold contacts on the upper surface. With this device, Bardeen and Brattain were able to demonstrate current gain.

¹ Bell Telephone Labs, in Murray Hill, New Jersey. The physical layout, performance, and physics of the point contact transistor is described in W. D. Bevitt’s book *Transistors Handbook*, chapter 3, and in Lo, et al., *Transistor Electronics*, chapter 1.14.

² It was John Bardeen who coined the terms *emitter*, *base*, and *collector*. For instance, in an NPN transistor, the emitter emits electrons and the collector collects them. The “base” was the base of a germanium crystal. See, e.g., W. F. Brinkman, et al., “A History of the Invention of the Transistor and Where It Will Lead Us.” An excellent reference describing the invention and implementation of the point-contact transistor is given in John Bardeen’s 1956 Nobel lecture.

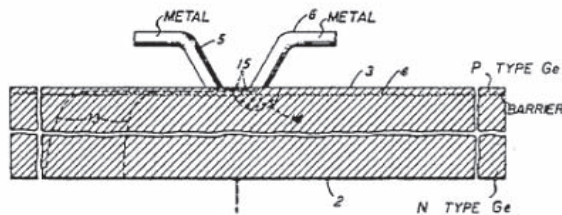


Fig. 2. Schematic of the point-contact transistor [3].

Figure 4-1: Point contact transistor from US Patent # 2,524,035 (June 17, 1948).

Later on, Shockley realized that a device with two semiconductor junctions could also demonstrate current gain. This led to the development of his transistor (**Figure 4-2**), which was the first bipolar junction transistor (BJT) and arguably the first practical transistor design. Shockley patented his device as well as a number of practical amplifying circuits. He had a remarkable ability to simplify and describe the fundamental physics governing carrier motion in diodes and transistors. The results of his pioneering work, coupled with work done by colleagues, led to the semiconductor revolution of the second half of the 20th century.

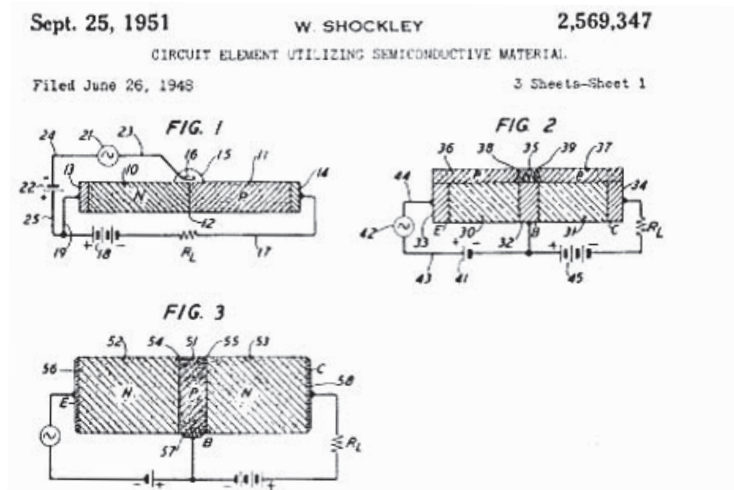
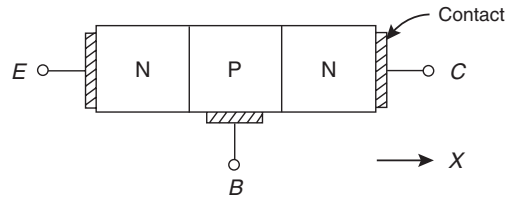


Figure 4-2: Shockley's transistor from US Patent # 2,569,346 (Sept. 25, 1951).

Basic NPN Transistor

We'll discuss some of the physics of the bipolar transistor by considering a basic one-dimensional structure. The structure of this basic NPN transistor (not to scale) is shown in **Figure 4-3**. It is comprised of an N-type emitter region, a P-type base, and an N-type collector region. Metallic, high-conductivity contacts are connected to an external circuit. This picture is cartoonish in that it doesn't show how the transistor is fabricated. However, it provides a useful framework through which we can analyze the operation. We'll assume a one-dimensional geometry (in x) throughout this chapter.

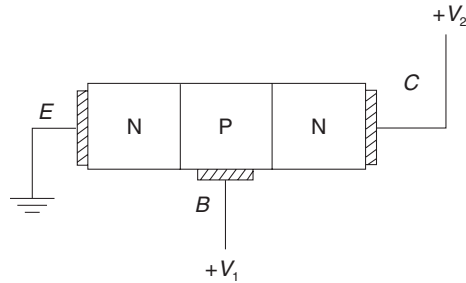
Figure 4-3: Basic NPN transistor showing emitter (E), base (B) and collector (C).



Now, let's talk through the operation of this BJT strike in so-called *thermal equilibrium*, when there are no connections to the device and hence no net current flow. From our work in the previous chapter, we know that there are drift and diffusion components working inside the semiconductor material. If there is no net current flow from any of the terminals, we know that the drift and diffusion currents exactly balance each other out, leaving no currents at the terminals. Furthermore, the hole drift and hole diffusion currents balance each other out, and the electron drift balances the electron diffusion currents.

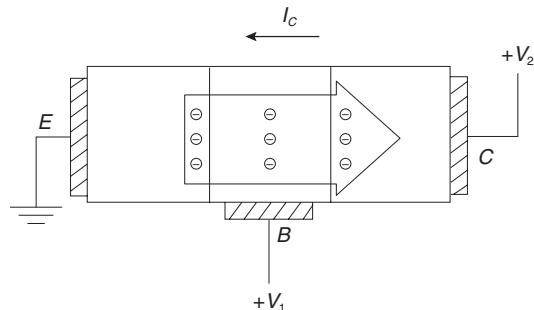
Let's first focus our efforts on the base-emitter junction. We know that an electric-field barrier to current flow exists at this PN junction. Now, if we connect the NPN transistor as shown in **Figure 4-4**, we can add a positive voltage to the base-emitter junction. The effect of this forward-bias voltage is to reduce the electric field barrier.

Figure 4-4: Basic NPN transistor connected with the base-emitter junction forward-biased.



Now that we've forward-biased the base-emitter junction and reduced the electric field barrier, there is less of an impediment for electrons to be injected from the emitter region to the collector region. As shown in **Figure 4-5**, electrons are now injected from the emitter, travel through the base region, and are collected by the collector. Since each electron carries a negative charge, our convention for the direction of current flow in an NPN transistor is as shown.

Figure 4-5: Basic NPN transistor connected with the base-emitter junction forward-biased, and the collector-base junction reverse-biased.



Let's figure out the cause of base current. We know that a forward-biased transistor requires a small but finite base current I_B to support the collector current I_C . The transistor current gain β_F is defined as:

$$\beta_F = \frac{I_C}{I_B} \quad [4-1]$$

What is the cause of the base current? The individual base current components are shown in **Figure 4-6**.³ Total base current is made up of three individual components, labeled I_{B1} , I_{B2} and I_{B3} .

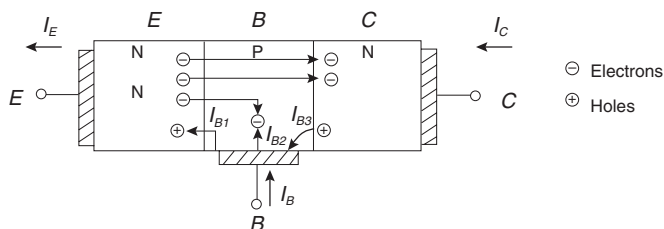
The emitter injects lots of electrons into the base region, as shown. Most of these electrons diffuse successfully across the base (as the base is very thin⁴). However, a small fraction of these injected electrons recombine with holes. A base current I_{B2} is needed to resupply the hole concentration in the base.

A second current component (I_{B1}) is comprised of holes that are injected from the base into the emitter.

There are holes created in the collector N-region by thermal generation. Some of these holes cross through the reverse-biased collector-base junction into the base. A third base current component I_{B3} is comprised of these holes that are injected across the reverse-biased base collector junction. This current component is the reverse leakage current of the reverse-biased collector junction.

The net base current I_B is the sum of these three individual components, and in a well-designed transistor this base current is a small fraction of the total collector current. The DC current gain β_F of a signal transistor can be typically 100 or greater.⁵

Figure 4-6: NPN transistor, forward-active region, showing base current components I_{B1} , I_{B2} and I_{B3} .



- ³ An excellent description of the PNP transistor, showing a similar construction, is given in Gerold Neudeck's *The Bipolar Junction Transistor*, Modular Series on Solid State Devices, volume 3, pp. 8–10.
- ⁴ In fact, transistor action depends on the base being very thin, so that most of the injected electrons will make it across the base. If the base width was very wide, the transistor would behave like back-to-back diodes, and all of the injected electrons would recombine in the base.
- ⁵ This holds true for a typical signal transistor. See, e.g., the 2N3904 transistor where at a bias level of $I_C = 10$ milliamps and $V_{CE} = 1$ volt, the DC beta β_F (called h_{FE} on the datasheet) shows a range of 100~300. Power transistors, on the other hand, often have β_F less than 10 or so, due to other design trade-offs. Current gain β_F also varies with collector current bias level. At very low and very high collector currents the β_F falls off.

A PNP transistor operating in the forward-active region is shown in **Figure 4-7**, with the base current components shown. The roles of electrons and holes are interchanged in the PNP transistor as compared to the NPN.

The bulk of the collector current is comprised of holes that are injected from the emitter into the base and which diffuse across the base. Base current I_{B2} is comprised of electrons flowing into the base that recombine with holes that are diffusing across the base. Base current I_{B1} consists of electrons that are injected from the base into the emitter. Base current I_{B3} is comprised of generated electrons in the collector that travel from the collector to the base.

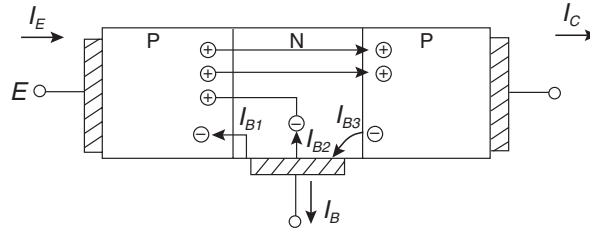


Figure 4-7: Basic PNP transistor, operating in the forward-active region, showing current components of the total base current I_B .

An NPN transistor, showing terminal variables, is shown in **Figure 4-8**. In an NPN transistor operating in the forward active region, the terminal currents (as defined in this diagram) are all positive. The base-emitter junction is forward-biased with $V_{BE} \sim 0.7V$ and the base-collector junction is reverse biased with $V_{BC} \leq 0$.

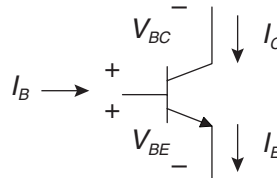


Figure 4-8: NPN transistor showing terminal variables.⁶

Transistor Models in Different Operating Regions

When we put an NPN transistor on a curve tracer (**Figure 4-9**), we find that the transistor has four regions of operation:

- In the cutoff region, $I_B = I_C = 0$ and $V_{BE} \ll 0.7V$. The transistor is essentially OFF.
- In the forward-active or linear region, $I_C = \beta_F I_B$ and $V_{BE} = 0.7V$ or so.
- In saturation, V_{CE} is small, and both the base-emitter and base-collector junctions are forward biased.

⁶ Some texts show I_E traveling in the opposite direction.

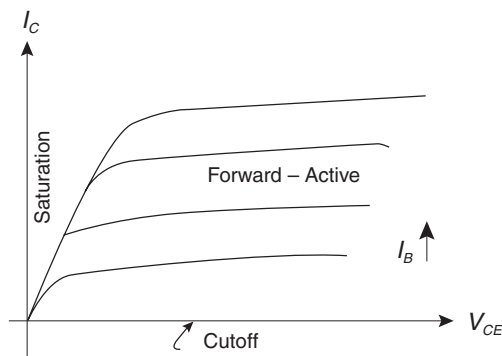
- In the reverse-active region (not shown on the diagram), the roles of collector and emitter are interchanged. Essentially, the transistor is being used “backwards.”⁷

In the forward-active region,⁸ the equations describing collector and base currents are:

$$I_C = I_S \left(e^{\frac{qV_{BE}}{kT}} - 1 \right) \quad [4-2]$$

$$I_B = \frac{I_C}{\beta_F}$$

Figure 4-9: NPN transistor on a curve tracer. There are three regions of operation in this quadrant: forward-active region, cutoff and saturation. Not shown is the reverse-active region.



In normal operation in transistor amplifiers, the transistor is used in the forward-active region, where:

$$I_C \approx I_S e^{\frac{qV_{BE}}{kT}} \quad [4-3]$$

In order to see how the various regions of operation work, let’s do a thought experiment with a simple resistively loaded NPN transistor (**Figure 4-10a**). We’ll initially assume that the base voltage V_B is zero volts. At $t = 0$, the base voltage starts ramping up slowly as shown. In this thought experiment, we’ll assume that the ramp is slow enough that bandwidth limitations inside the transistor don’t come into play.⁹

In **Figure 4-10b** are the transistor characteristic curves with the load-line associated with the 1-k Ω resistor superimposed. Before $t = 0$, the operating point is shown at location A in the diagram. The transistor is cutoff, with zero collector current and with $V_{CE} = 12$ volts as shown.

⁷ Note that if you interchange the collector and emitter leads in a conventional transistor circuit, the device will operate as a transistor. However, the β_F will be low and the speed will be low as well. Transistors are optimized to operate in the forward-active region.

⁸ Remember that the reverse saturation current I_S is the leakage current drawn by the ideal PN junction when it is reverse biased.

⁹ In this case we’ll use a ramp that transitions from 0V to 1V in 1 second.

Some time after $t = 0$, when the base voltage V_B has risen to ~ 0.4 volts or so,¹⁰ the transistor starts turning on and begins to conduct collector current. During this region of operation, the transistor is in the “forward-active” region, labeled B in the diagram. When the collector current rises, the transistor collector-emitter voltage falls, due to voltage drop across the $1\text{-k}\Omega$ resistor.

As the base voltage continues to rise, the collector voltage falls, until we reach region C in the curves. The transistor is now saturated. When the transistor is saturated, there is a relatively low value (less than 0.25 volts or so) across the collector-emitter. As the base voltage increases, we go further and further into saturation, when the maximum collector current of approximately 12 milliamps is reached.

This switching profile is shown in **Figure 4-10c**, where we plot base voltage and collector current. Note, as expected, that the transistor begins turning on at $V_B \sim 0.4$ volts or so, and that the maximum collector current is approximately 12 milliamps.

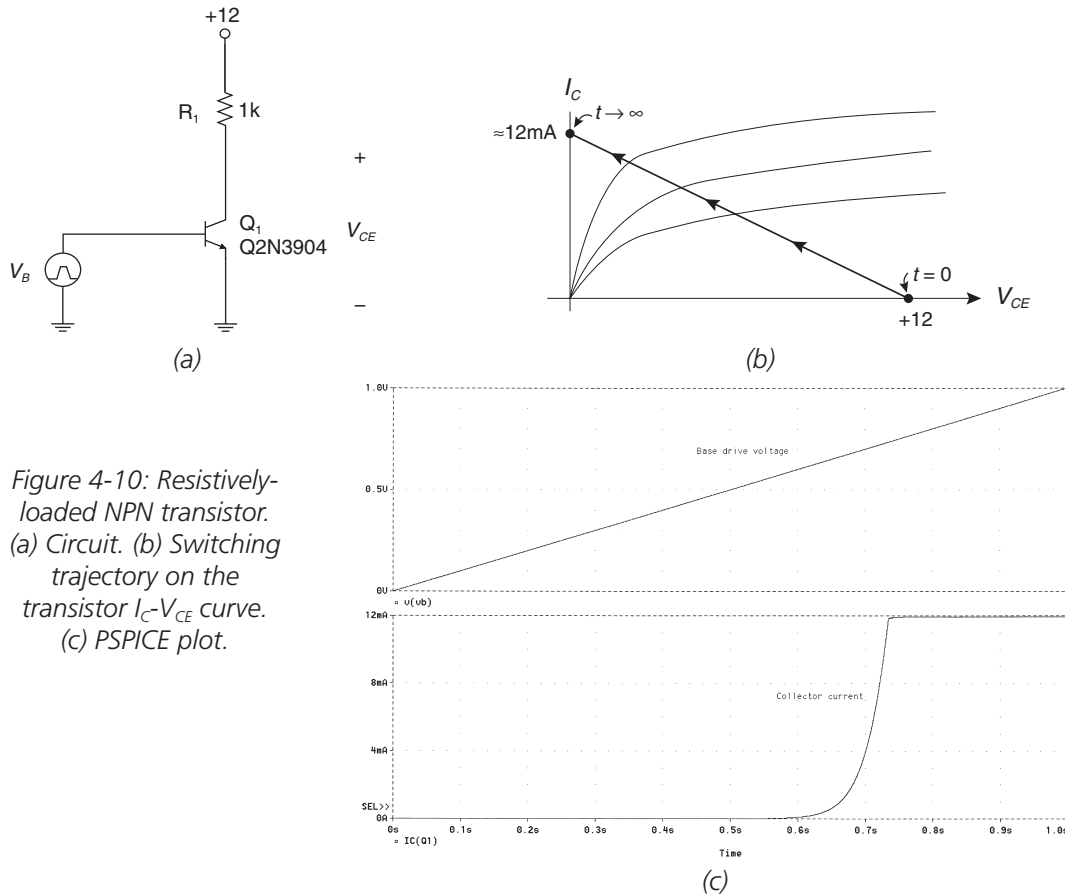


Figure 4-10: Resistively-loaded NPN transistor.
(a) Circuit. (b) Switching trajectory on the transistor I_C - V_{CE} curve.
(c) PSPICE plot.

¹⁰ As a rule-of-thumb we normally assume that a diode forward voltage is approximately 0.7V; a transistor when it's biased in the forward active region has a base-emitter drop of approximately 0.7V. This is the nature of a forward-biased PN junction. However, the transistor does begin to turn on at a somewhat lower voltage. We'll estimate the beginning of the turn-on at $V_{BE} \sim 0.4\text{V}$.

Low-Frequency Incremental Bipolar Transistor Model

The low-frequency incremental model follows directly from the large-signal model. We'll linearize about the operating point to get an approximately linear model that is valid for small variations in terminal voltages and currents.

Let's assume that we have a resistively loaded NPN transistor properly biased so that the transistor operates in the middle of the forward active region.¹¹ First, we'll assume that all transistor terminal variables have a DC part and a small-signal varying part:

$$\begin{aligned} i_C &= I_C + i_c \\ i_B &= I_B + i_b \\ v_{BE} &= V_{BE} + v_{be} \\ &\text{etc.....} \end{aligned} \tag{4-4}$$

where the term on the left is the total variable, and the terms on the right are the DC (or bias point) and small-signal part, respectively. In **Figure 4-11** we have illustrated this for the collector current.

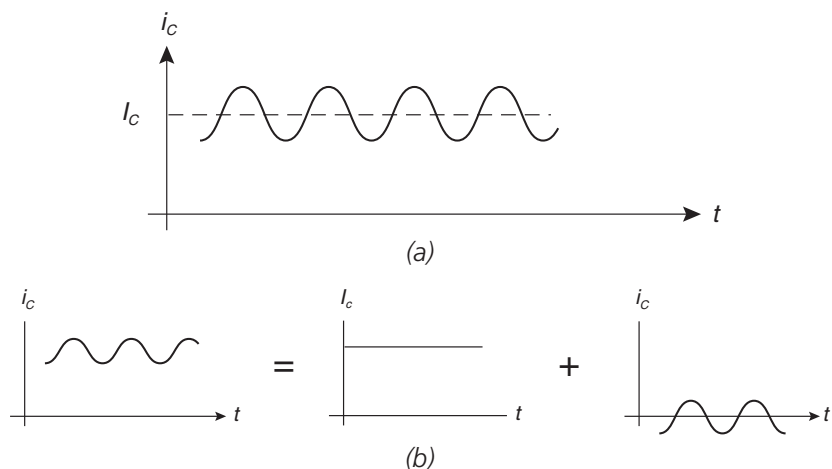


Figure 4-11: Diagram illustrating total collector current (i_C), bias level of collector current (I_C), and incremental part of collector current (i_c). (a) Total collector current $i_C(t)$. (b) Total collector current (i_C) broken into individual components of the DC bias level (I_C) and incremental variation (i_c).

In **Figure 4-12a**, we show a transistor circuit biased with a base battery V_{BIAS} and an input signal v_{sig} which has a peak-to-peak amplitude of 1 millivolt at a frequency of 1 kHz. The level of V_{BIAS} is approximately 0.7 volts,¹² but let's assume that we can adjust it so that the

¹¹ Initially, at least, we'll ignore the practical problem of how to bias this transistor at this operating point so that the bias point is stable with respect to temperature and component variations.

¹² The actual value of V_{BIAS} used in the SPICE simulation is 706.5 millivolts, which sets up a bias current of approximately 5 milliamps.

transistor collector current is 5 milliamps. This will bias the transistor in the middle of the forward-active region as shown in **Figure 4-12b**. When we wiggle v_{sig} the collector current varies sinusoidally as shown in **Figure 4-12c**.

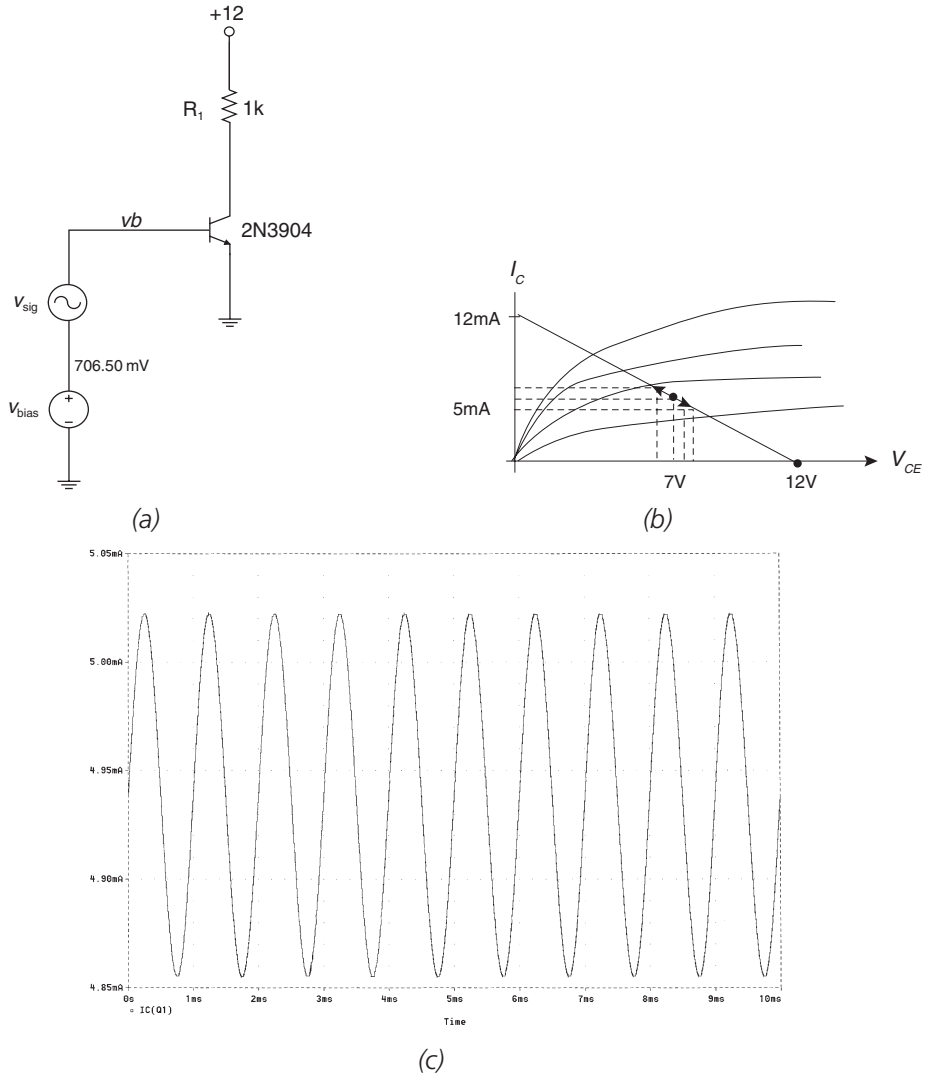


Figure 4-12: Resistively loaded NPN transistor biased in the forward-active region. (a) Circuit. (b) Operating point shown on the transistor I_C - V_{CE} curve. (c) PSpice plots of above circuit showing AC output of the collector current riding on a DC bias level of ~ 4.95 mA.

For the collector current with the transistor biased in the forward active region, we find again the relationship between base-emitter voltage and collector current:

$$I_C + i_c \approx I_S e^{\frac{q(V_{BE} + v_{be})}{kT}} = I_C e^{\frac{qv_{be}}{kT}} \quad [4-5]$$

Chapter 4

We can approximately ignore the “–1” term in the exponential transistor voltage-current equation if we assume that the operating point value of V_{BE} is much greater than kT/q . If we restrict the *small-signal variation* of v_{be} to be small compared to kT/q (26 mV at room temperature) we can use the approximation $e^x \approx 1 + x$ to get:

$$I_C + i_c \approx I_S e^{\frac{q(V_{BE} + v_{be})}{kT}} \approx I_C \left(1 + \frac{qv_{be}}{kT} \right) \quad [4-6]$$

Solving for the small-signal component i_c of the collector current results in:

$$i_c \approx I_C \left(\frac{qv_{be}}{kT} \right) = g_m v_{be} \quad [4-7]$$

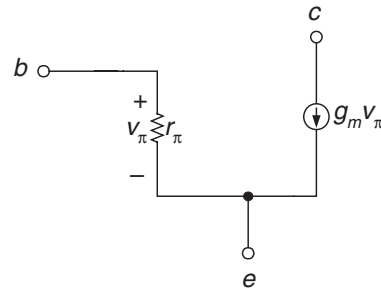
We have defined a proportionality constant g_m , which relates *small signal variation* in collector current (i_c) to *small signal variation* in v_{be} . This term g_m is called the *transconductance* of the transistor and has units of amps per volt.

Now, let’s focus on small-signal variations in base current. We know that a small-signal variation in collector current means a proportional small-signal variation in base current. The base current can be described, using this reasoning, as:

$$i_b \approx \frac{i_c}{h_{fe}} = \frac{g_m}{h_{fe}} v_{be} = \frac{v_{be}}{r_\pi} \quad [4-8]$$

where h_{fe} is the small-signal current gain¹³ of the transistor (often denoted h_{fe} on the transistor datasheet, and often denoted β_o in textbooks). The model in **Figure 4-13** captures the functional dependence of the collector current and base current equations. This so-called “hybrid-pi” model is useful for finding midband gain of a transistor amplifier, and we’ll use it in subsequent chapters extensively.

Figure 4-13. Bipolar transistor low-frequency incremental model, showing base (b), collector (c) and emitter (e) terminals.



To summarize, the low-frequency incremental model parameters for the bipolar transistor are:

$$g_m = \frac{|I_C|}{kT/q} = \frac{|I_C|}{V_{TH}} \quad [4-9]$$

$$r_\pi = \frac{h_{fe}}{g_m}$$

¹³ Don’t get small signal beta (β_o or h_{fe}) confused with DC or operating point beta (β_F or h_{FE}).

High-Frequency Incremental Model

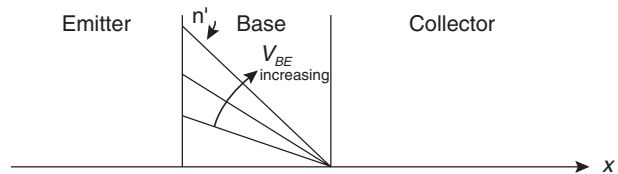
We know that the previous low-frequency model is incomplete, because it does not have any bandwidth-limiting mechanisms. Let's now consider some of the sources of bandwidth limitation in bipolar transistors.

First, we know that a junction has a capacitance associated with it, as we showed in the diode discussion of the previous chapter. Therefore, we expect that there will be junction capacitances in the bipolar transistor associated with the base-emitter and base-collector junctions. These junction capacitances¹⁴ show a functional dependence on junction voltage as shown in the previous chapter.

Secondly, when the transistor is biased in the forward-active region, there are charges stored in the base (**Figure 4-14**). As v_{BE} varies up and down, the concentration of extra electrons (n') stored in the base varies up and down. In this formulation, we'll assume that the variation in v_{BE} occurs slowly enough so that the concentration of n' in the base can be modeled as a series of static, triangular distributions. This is the so-called “quasi-static” approximation and will be used in later chapters when we discuss transistor large-signal switching.

This stored base charge can be modeled as a capacitance that depends on the bias level of the transistor. In the case of the NPN transistor, as V_{BE} increases, the excess minority carrier concentration stored in the base increases as well, as shown.

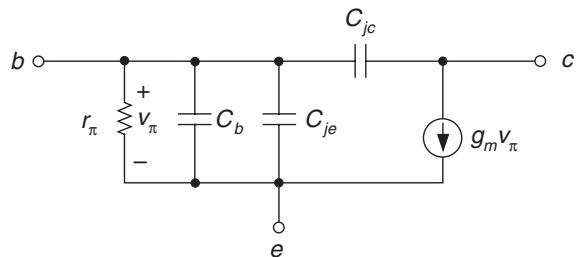
Figure 4-14: Charge stored in the base region of a transistor in the forward-active region. As transistor base-emitter voltage V_{BE} increases, the extra electrons in the base n' increase as shown.



A circuit model showing the various charge storage mechanisms in a BJT is shown in **Figure 4-15**. There are the following capacitances in the model:

- C_{je} : Base-emitter depletion capacitance
- C_b : Base diffusion capacitance
- C_{jc} : Base-collector depletion capacitance

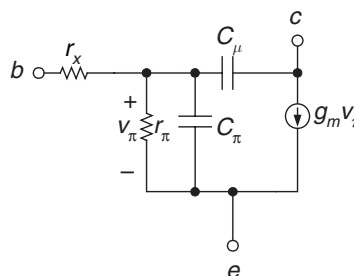
Figure 4-15: Transistor model showing charge storage components C_b (base diffusion capacitance), C_{je} (base-emitter junction capacitance) and C_{jc} (base-collector junction capacitance).



¹⁴ The junction capacitances are also called *space charge* capacitances or *depletion* capacitances.

To this model, we make several modifications (**Figure 4-16**). First, we've combined C_{je} and C_b into a single capacitance which we'll call C_π . Secondly, let's rename C_{jc} and call it C_μ . Lastly, let's add the base spreading resistance r_x . This models the resistance between the ohmic base contact and is due in part to two-dimensional base current flow effects. It is important to include r_x in the model because the transistor capacitances C_π and C_μ must be charged through it. If you omit r_x , in certain circuit topologies your model will be overly optimistic as to the bandwidth.

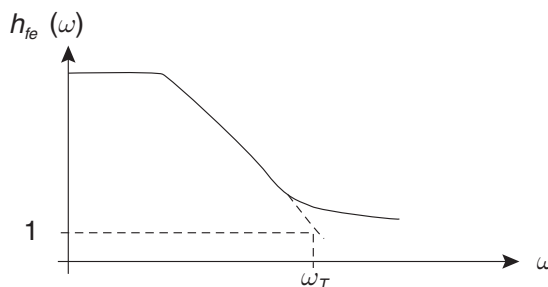
Figure 4-16: High frequency incremental model. In this model we've combined C_{je} and C_b into a single capacitor C_π , have renamed C_{jc} as C_μ , and have added base resistance r_x .



Now, how do we determine C_π and C_μ from the datasheet? C_μ is relatively easy. We know that C_μ is just the base-collector junction capacitance. First, we need to determine the operating point value of the base-collector junction voltage. Then, we can just read the junction capacitance off the datasheet at the given base-collector bias voltage.

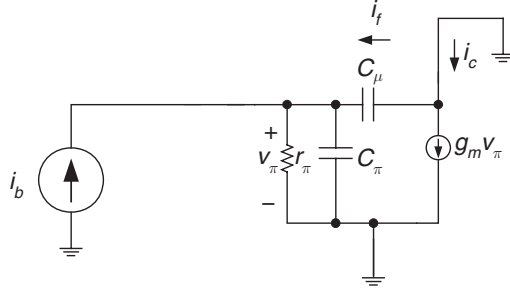
In order to find C_π , we need to make use of some information from the datasheet. If you look at the datasheet, there is a number listed, sometimes called the *transition frequency* or *current gain-bandwidth product*. If we look at a plot of transistor current gain vs. frequency we'll see a plot something like **Figure 4-17**. The current gain-bandwidth number is the frequency at which extrapolated small-signal current gain reaches unity.

Figure 4-17: Plot of incremental current gain $h_{fe}(\omega)$ vs. frequency. At frequency ω_T , the extrapolated curve reaches a current gain of 1.



We can use the simple circuit of **Figure 4-18** to help us figure out a methodology for finding C_π . First, for simplicity, we have ignored the effects of r_x . Let's solve for incremental collector i_c when the base is driven by an incremental base current i_b .

Figure 4-18: Transistor incremental model for finding current gain-bandwidth product f_T .



Since the right-hand side of C_μ is grounded, the voltage v_π is simply:

$$v_\pi = i_b \frac{r_\pi}{r_\pi (C_\pi + C_\mu)s + 1} \quad [4-10]$$

The collector current i_c is:

$$\begin{aligned} i_c &= g_m v_\pi + i_f \\ &= g_m i_b \frac{r_\pi}{r_\pi (C_\pi + C_\mu)s + 1} - i_b \frac{r_\pi C_\mu s}{r_\pi (C_\pi + C_\mu)s + 1} \\ &= \frac{(h_{fe} - r_\pi C_\mu s) i_b}{r_\pi (C_\pi + C_\mu)s + 1} \end{aligned} \quad [4-11]$$

Note that the feedback through C_μ gives a zero in the right-half plane at $\omega_z = +1/r_\pi C_\mu$, which is at a frequency higher than ω_T . Since the zero frequency is higher than the frequency range that we are interested in,¹⁵ we'll ignore it, approximating the transfer function by:

$$\frac{i_c}{i_b} \approx \frac{h_{fe}}{r_\pi (C_\pi + C_\mu)s + 1} \quad [4-12]$$

For frequencies much higher than the breakpoint, and making use of the fact that $h_{fe} = \beta_o = g_m r_\pi$ results in:

$$\frac{i_c}{i_b} \approx \frac{g_m r_\pi}{r_\pi (C_\pi + C_\mu)s} = \frac{g_m}{(C_\pi + C_\mu)s} \quad [4-13]$$

Therefore, the magnitude where this gain drops to 1 is:

$$\omega_T \approx \frac{g_m}{(C_\pi + C_\mu)} \quad [4-14]$$

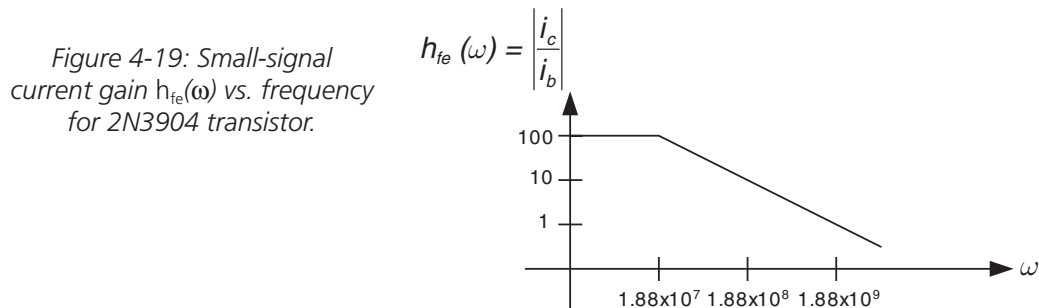
¹⁵ This frequency is higher than the frequency where the transistor current gain drops to 1. Therefore, in the interest of mathematical simplicity, we simply throw it away.

Therefore, our recipe for finding C_π and C_μ from the datasheet and bias information is:

$$C_\mu \approx C_{jc} \text{ (found on datasheet at } V_{BC} \text{ bias)}$$

$$C_\pi = \frac{g_m}{\omega_T} - C_\mu \quad [4-15]$$

For the 2N3904 transistor with $I_C = 10$ milliamps, $h_{fe} \approx 100$ and $\omega_T = 2\pi \times 300 \times 10^6$ radians/second = 1.885×10^9 rad/sec. Transistor h_{fe} begins dropping at $\omega = 1.885 \times 10^7$ radians per second.



Reading a Transistor Datasheet

We'll now go through the exercise of reading sections of a commercially available signal transistor datasheet (for the 2N3904¹⁶) and extracting important information. The 2N3904 is an inexpensive signal transistor with an approximate f_T of 300 MHz available in a plastic TO-92 package (Figure 4-20) with pins 1, 2 and 3 being the emitter, base and collector, respectively.

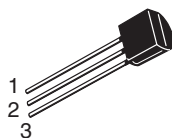


Figure 4-20: TO-92 plastic package.

Large signal parameters (β_F , $V_{CE,SAT}$)

The large-signal transistor parameter that we'll consider first is DC current gain, or β_F (Figure 4-21). On the datasheet,¹⁷ this is often called " h_{FE} ." The charts indicate that the 2N3904 has a h_{FE} that peaks for intermediate levels of collector current. For instance, at 10 mA collec-

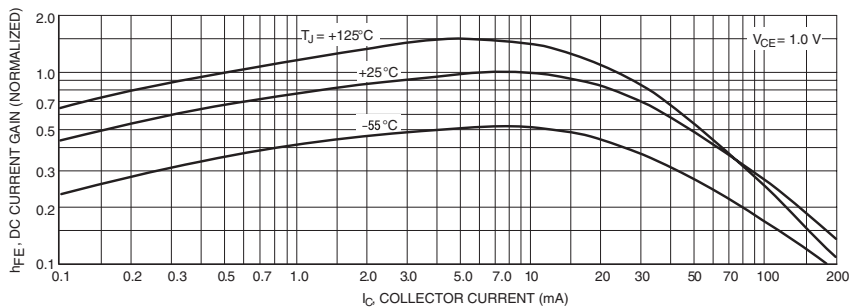
¹⁶ I'll use the datasheet from On Semiconductor (www.onsemi.com) but this same device is made by Philips, Fairchild, Vishay, National Semiconductor, and other companies. We'll extract information from the datasheet in this chapter germane to the topics of transistor biasing and small-signal modeling. In a later chapter on the charge control model and transistor switching, we'll extract further information. Reprinted with permission of On Semiconductor.

¹⁷ For historical reasons, DC beta is usually termed β_F in textbooks and journal articles, while in datasheets it's h_{FE} .

tor current, a typical number for h_{FE} is 300, with a minimum of 100. **Figure 4-21b** also shows this variation of h_{FE} with collector current, as well as the variation with temperature. Note that h_{FE} increases as temperature increases, at least for modest collector current. h_{FE} also drops at low collector currents due to recombination of carriers in the base-emitter depletion region. At high collector currents, h_{FE} falls off primarily due to emitter current crowding.

ON CHARACTERISTICS		h_{FE}			
DC Current Gain (Note 2)					
(I _C = 0.1 mA, V _{CE} = 1.0 Vdc)	2N3903	20	—		
	2N3904	40	—		
(I _C = 1.0 mA, V _{CE} = 1.0 Vdc)	2N3903	35	—		
	2N3904	70	—		
(I _C = 10 mA, V _{CE} = 1.0 Vdc)	2N3903	50	150		
	2N3904	100	300		
(I _C = 50 mA, V _{CE} = 1.0 Vdc)	2N3903	30	—		
	2N3904	60	—		
(I _C = 100 mA, V _{CE} = 1.0 Vdc)	2N3903	15	—		
	2N3904	30	—		

(a)



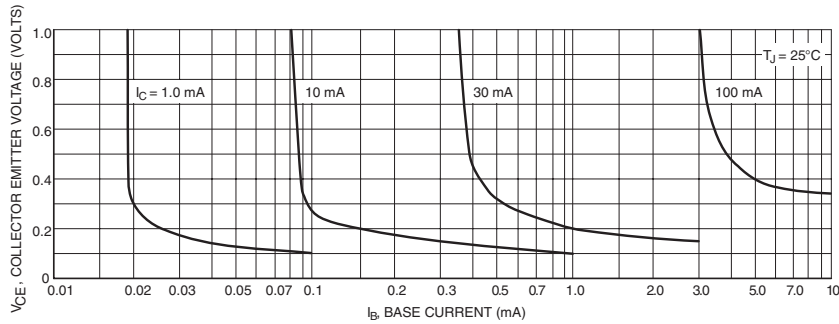
(b)

Figure 4-21: Datasheet information for 2N3904 for DC current gain h_{FE} (also called β_F).
(a) Datasheet values. (b) Curves. Note the variation in DC current gain with temperature, as well as variation with bias collector current value.
(Reprinted with permission of On Semiconductor.)

The value of saturation voltage for the 2N3904 is shown in **Figure 4-22**. Note that in hard saturation the value of $V_{CE,SAT}$ can be as low as 0.1V or less, depending on how hard the base is driven with base current.

Collector–Emitter Saturation Voltage (Note 2) ($I_C = 10 \text{ mAdc}$, $I_B = 1.0 \text{ mAdc}$) ($I_C = 50 \text{ mAdc}$, $I_B = 5.0 \text{ mAdc}$)	$V_{CE(sat)}$	– –	0.2 0.3	V_{dc}
---	---------------	--------	------------	----------

(a)



(b)

Figure 4-22: Datasheet information for 2N3904 on saturation characteristics ($V_{CE,SAT}$). Note that as you increase the base current overdrive, the saturation voltage decreases.

Small signal parameters (h_{fe} , C_{μ} , C_{π} and r_x)

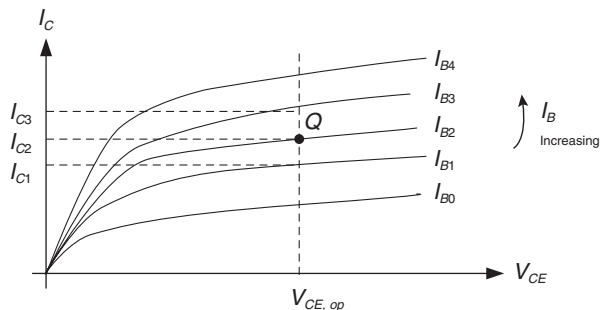
Small-signal beta (sometimes called β_o in textbooks, and often h_{fe} on datasheets) is not the same as large-signal β_F . Large-signal β_F tells you what the bias value of the base current (I_B) will be if you know the bias value of the collector current (I_C). Small-signal β_o tells you what the variation will be about those operating points, or:

$$\beta_o = h_{fe} = \left. \frac{\partial i_c}{\partial i_b} \right|_{I_C} \quad [4-16]$$

If you put an NPN transistor on a curve tracer, you will get a plot something like **Figure 4-23**. Let's assume that the point Q is the operating point where we will be using the transistor in our circuit. Using the curve tracer, we can find the small-signal beta h_{fe} near this operating point as:

$$h_{fe} = \left. \frac{\Delta i_c}{\Delta i_b} \right|_{I_C} = \frac{I_{C3} - I_{C1}}{I_{B3} - I_{B1}} \quad [4-17]$$

Figure 4-23: Curve tracer method for finding small-signal current gain h_{fe} measured around your operating point Q.



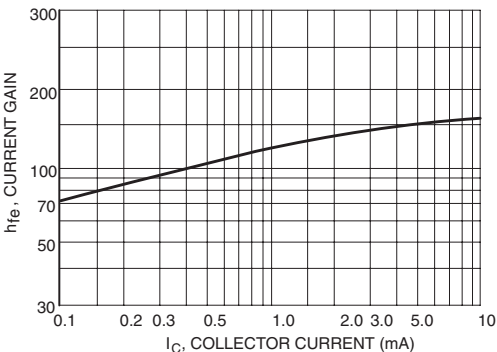
Datasheet information for h_{fe} for the 2N3904 is shown in **Figure 4-24**. If we know the bias level of the collector current we can find transconductance g_m . If we know g_m and h_{fe} , we can find r_π :

$$r_\pi = \frac{h_{fe}}{g_m} \tag{4-18}$$

Small-Signal Current Gain ($I_C = 1.0\text{ mA}$, $V_{CE} = 10\text{ V}$, $f = 1.0\text{ kHz}$)	2N3903 2N3904	h_{fe}	50 100	200 400	—
--	------------------	----------	-----------	------------	---

(a)

Figure 4-24: Datasheet information for 2N3904 on small-signal β_o (or h_{fe}).
(a) Datasheet table. (b) Curve.
Note that h_{fe} varies with the bias level of the collector current.



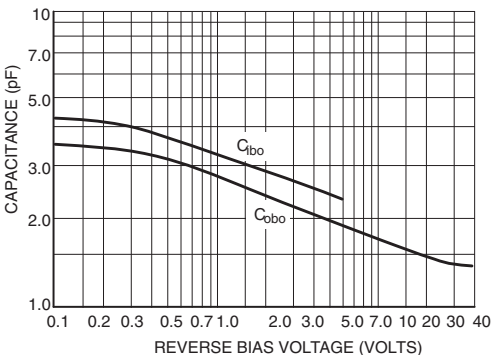
(b)

Depletion capacitance information for the 2N3904 is shown in **Figure 4-25**. The curve C_{obo} is the base-collector junction capacitance, and C_{ibo} is the base-emitter junction capacitance.

Output Capacitance ($V_{CB} = 5.0\text{ V}$, $I_E = 0$, $f = 1.0\text{ MHz}$)	C_{obo}	—	4.0	pF
Input Capacitance ($V_{EB} = 0.5\text{ V}$, $I_C = 0$, $f = 1.0\text{ MHz}$)	C_{ibo}	—	8.0	pF

(a)

Figure 4-25: Datasheet information for 2N3904 on junction capacitances.
(a) Table value. (b) Curve. Note that as the reverse bias voltage on the junction increases, the junction capacitances decrease.



(b)

Chapter 4

Information on base spreading resistance r_x is usually not found directly on transistor data-sheets. Some transistors do provide other information, such as the *collector base time constant* shown in **Figure 4-26a** for the 2N2222 transistor. We see in this example that the time constant may be used to find r_x , as follows. The collector base time constant is specified to be 150 picoseconds:

$$r_b C_c \approx r_x C_\mu = \tau_{cb} = 150 \times 10^{-12} \quad [4-19]$$

We note that this value of collector base time constant is specified at a collector-base reverse bias voltage $V_{CB} = 20$ volts. We now look at the chart of C_{cb} vs. frequency (**Figure 4-26a**) and find that at this reverse voltage $C_{cb} = C_\mu \approx 3.4$ picofarads. Next, we can estimate base spreading resistance as:

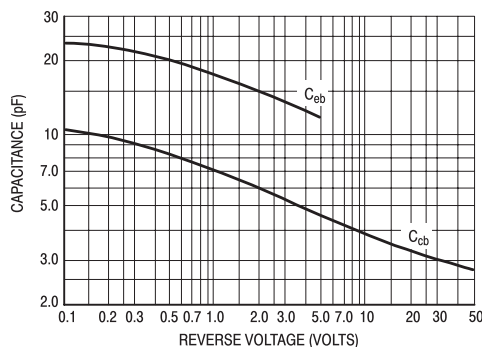
$$r_x \approx \frac{\tau_{cb}}{C_\mu} = \frac{150 \times 10^{-12}}{3.4 \times 10^{-12}} \approx 44 \Omega \quad [4-20]$$

Note that this value of r_x is approximate at an emitter current (and hence a collector current) of 20 milliamps. The value varies significantly with collector current due to current crowding. At lower collector currents, the value of r_x will be somewhat higher, but the functional dependence depends on details of the device fabrication.

Collector Base Time Constant ($I_E = 20$ mAdc, $V_{CB} = 20$ Vdc, $f = 31.8$ MHz)	$r_b C_c$	—	150	ps
---	-----------	---	-----	----

(a)

Figure 4-26: Datasheet information for 2N2222¹⁸ on base-collector time constant. (a) Time constant found on the datasheet. (b) Information on junction capacitances.



(b)

Summarizing, in order to find appropriate small-signal models, this is a good design process:

- Find operating point values for collector current I_C and collector-base voltage V_{CE} . If the transistor is in the forward-active region, move on to find small-signal parameters.
- Find transistor transconductance g_m .
- From the datasheet, find h_{fe} at your bias collector current.

¹⁸ Datasheet from On Semiconductor, www.onsemi.com, reprinted with permission of On Semiconductor.

- Calculate r_π .
- From the datasheet, find output capacitance C_{obo} (which is the base-collector junction capacitance¹⁹). Small-signal capacitance $C_\mu = C_{obo}$.
- From the datasheet, find f_T (or ω_T). Note that f_T is collector-current-dependent.
- Calculate C_π .
- In the absence of any other additional information, make an educated guess as to the value of base spreading resistance r_x .

Limitations of Hybrid-Pi Model

A fundamental limitation of the hybrid-pi model is that it is valid for frequencies much less than the f_T of the transistor. Let's see what this means for our 2N3904 example.

The datasheet value for f_T of the 2N3904 is 300 MHz. (Note that this value of f_T varies with collector current level; unfortunately this current variation is not to be found on this particular datasheet.) The variation in f_T with collector current level for a typical transistor is shown in **Figure 4-27**. At low values of collector current, junction capacitances C_{je} and C_{jc} dominate and f_T is approximately linearly increasing with collector current. At very high collector current, f_T drops off due to high-level injection effects.²⁰

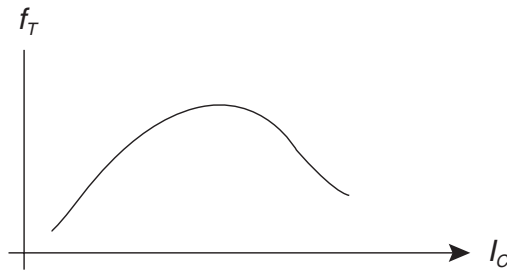


Figure 4-27: Variation in f_T with collector current for hypothetical small-signal transistor.

¹⁹ Some manufacturers use different names. For instance, in the 2N2222 datasheet, the collector base junction capacitance is named " C_{cb} ."

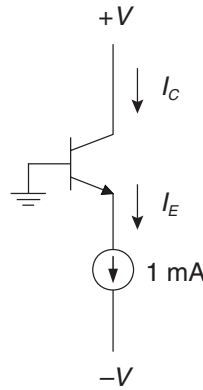
²⁰ See, e.g., John Choma, Jr., "A Curve-Fitted Circuits Model for Bipolar f_T Roll-Off at High Injection Levels," *IEEE Journal of Solid-State Circuits*, April 1976, pp. 346–348.

Chapter 4 Problems

Problem 4.1

For the transistor circuit in **Figure 4-28**, find transistor collector current and small-signal parameters g_m and r_π , and draw the low-frequency small-signal model. Assume that large-signal current gain $\beta_F = 100$ and that small-signal current gain $h_{fe} = 150$, and that the base resistance $r_x = 200\Omega$. Assume room temperature operation of the transistor.

Figure 4-28: Transistor circuit for Problem 4.1



Problem 4.2

Find and sketch the low-frequency incremental model for a transistor with $I_c = 10\text{ mA}$ and $h_{fe} = 200$. Assume that the transistor operates at room temperature. Assume that the transistor base resistance $r_x = 0$.

Problem 4.3

A transistor has parameters $f_T = 500\text{ MHz}$, $C_{jc} = 1\text{ pF}$ and $h_{fe} = 100$. Find and sketch the high-frequency incremental model when this transistor is biased with a collector current of 1 milliamp. Assume that transistor base resistance $r_x = 50\Omega$.

Problem 4.4

Using datasheet information from the 2N3904 transistor, find the small-signal model assuming the transistor is biased at a collector current of 5 milliamps and a collector-base voltage of 5 volts.

Problem 4.5

Consider the transistor circuit shown in **Figure 4-29**. The base of the transistor is biased with bias current source I_B , and a small-signal AC current source is also connected, i_b . The resultant collector current has both a DC part, I_C , and a small-signal varying part i_c .

- Draw the small-signal model using the small-signal high-frequency model (i.e., including C_π and C_μ), but ignoring the base resistance r_x (i.e., assume r_x is very small).
- Calculate the small-signal current transfer ratio $i_c(s)/i_b(s)$ in terms of g_m , C_π and C_μ .
- Using a reasonable approximation and the result from (b), calculate the frequency ω_T (in radians/second) where the magnitude of the small-signal current gain $i_c(s)/i_b(s)$ drops to 1, in terms of g_m , C_π and C_μ . (*Hint: You should be able to do this graphically without significant calculations.*)
- For the 2N3904 transistor, biased at $I_C = 10$ mA, plot the magnitude of i_c/i_b using *minimum* numbers from the datasheet and on your plot denote the value of ω_T . Label all breakpoints, etc.

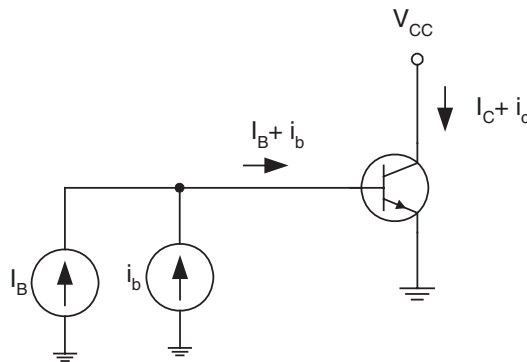


Figure 4-29: Transistor circuit for Problem 4.5.

References

- Bardeen, John, "Semiconductor Research Leading to the Point Contact Transistor," December 11, 1956, from *Nobel Lectures Physics, 1942-1962*, Elsevier Publishing, available from the Nobel e-museum at Web: <http://www.nobel.se/physics/laureates/1956/>
- Bardeen, J., and Brattain, W. H., "Physical Principles Involved in Transistor Action," *Physical Review*, vol. 75, 1949, p. 1208.
- Bevitt, W. D., *Transistors Handbook*, Prentice-Hall, 1956.
- Brinkman, W. F., et al., "A History of the Invention of the Transistor and Where It Will Lead Us."
- Choma, John, Jr., "A Curve-Fitted Circuits Model for Bipolar f_T Roll-Off at High Injection Levels," *IEEE Journal of Solid-State Circuits*, April 1976, pp. 346–348.
- Fair, R. B., "History of some early developments in ion-implantation technology leading to

- silicon transistor manufacturing,” *Proceedings of the IEEE*, vol. 86, no. 1, January 1998, pp. 111–137.
- Gough, Russell, “High-Frequency Transistor Modeling for Circuit Simulation,” *IEEE Journal of Solid-State Circuits*, vol. SC-17, no. 4, August 1982, pp. 666–670.
- Gray, P. E., DeWitt, D., Boothroyd, A. R., and Gibbons, J. F., *Physical Electronics and Circuit Models of Transistors*, Semiconductor Electronics Education Committee, volume 2, John Wiley, 1964.
- Gray, Paul E., and Searle, Campbell L., *Electronic Principles: Physics, Models and Circuits*, John Wiley, 1969.
- Gray, Paul R. and Meyer, Robert G., *Analysis and Design of Analog Integrated Circuits*, 2d edition, John Wiley, 1984.
- Greeneich, E. W., “An Appropriate Device Figure of Merit for Bipolar CML,” *IEEE Electron Device Letters*, vol. 12, 1991, p. 18.
- Hurkx, G. A. M., “The Relevance of f_T and f_{max} for the Speed of a Bipolar CE Amplifier Stage,” *IEEE Transactions on Electron Devices*, vol. 44, no. 5, May 1997, pp. 775–781.
- Lo, A. W., Endres, R., Zawals, J., Waldhauer, F., and Cheng, C., *Transistor Electronics*, Prentice-Hall, 1955.
- Muller, Richard S., and Kamins, Theodore I., *Device Electronics for Integrated Circuits*, 2d edition, John Wiley, 1986.
- Neudeck, Gerold W., *Modular Series on Solid State Devices, The Bipolar Junction Transistor*, Addison-Wesley, 1983.
- On Semiconductor website: <http://www.onsemi.com>
- Pritchard, R. L., “Transistor equivalent circuits,” *Proceedings of the IEEE*, vol. 86, no. 1, January 1998, pp. 150–162.
- Sansen, Willy, and Meyer, Robert, “Characterization and Measurement of the Base and Emitter Resistances of Bipolar Transistors,” *IEEE Journal of Solid-State Circuits*, vol. SC-7, no. 6, December 1972, pp. 492–498.
- Searle, C. L., Boothroyd, A. R., Angelo, E. J., Jr., Gray, P. E., and Pederson, D. O., *Elementary Circuit Properties of Transistors*, Semiconductor Electronics Education Committee, volume 3, John Wiley, 1964.
- Shockley, William, “Transistor Technology Evokes New Physics,” December 11, 1956, from *Nobel Lectures Physics, 1942–1962*, Elsevier Publishing, available from the Nobel e-museum at Web: <http://www.nobel.se/physics/laureates/1956/>
- Sze, S. M., *Physics of Semiconductor Devices*, 2d edition, John Wiley, 1981.
- Warner, Raymond M., “Microelectronics: Its Unusual Origin and Personality,” *IEEE Transactions on Electron Devices*, vol. 48, no. 11, November 2001, pp. 2457–2467.

Basic Bipolar Transistor Amplifiers and Biasing

In This Chapter

- In this chapter, we cover some basic transistor topologies, including the common-emitter amplifier, emitter-follower, common-base and differential amplifiers. The important issue of biasing—connection of the transistor so that it operates in the forward-active region—is also discussed in detail.

The Issue of Transistor Biasing

In order to achieve useful amplification from a transistor, the transistor is generally biased in the forward-active region (FAR). The FAR is the region of operation of the transistor where amplification can occur because the transistor provides current gain. Biasing is the process by which one sets the DC operating point of a transistor amplifier to a known and repeatable point in this forward-active region. If you do your biasing correctly, your amplifier output bias level will not drift significantly with time, temperature or component variations.

Let's consider the simple bias circuit shown in **Figure 5-1a**. The bias voltage $V_{BB} = 5\text{V}$ sets the base current as follows:

$$I_B = \frac{V_{BB} - V_{BE}}{R_B} \approx \frac{5 - 0.7}{28.6 \text{ k}\Omega} \approx 150 \text{ }\mu\text{A} \quad [5-1]$$

This initial calculation assumes that the base-emitter voltage $V_{BE} = 0.7\text{V}$, which is approximately correct if the transistor is biased in the forward-active region and the base-emitter junction is forward biased. Let's assume that this transistor has a DC current gain $\beta_F = 175$. This means that the collector current is:

$$I_C = \beta_F I_B = \beta_F \left(\frac{V_{BB} - V_{BE}}{R_B} \right) \approx 175 \times 150 \text{ }\mu\text{A} \approx 26.25 \text{ milliamps} \quad [5-2]$$

With a load resistor $R_L = 240\Omega$, there is a 6.3V drop across this load resistor and hence $V_{CE} = 5.7\text{V}$. The operating point, labeled Q for quiescent operating point, is shown in **Figure 5-1b**. The operating point is the intersection of the transistor characteristic curves with the load line imposed by the 240Ω resistor.

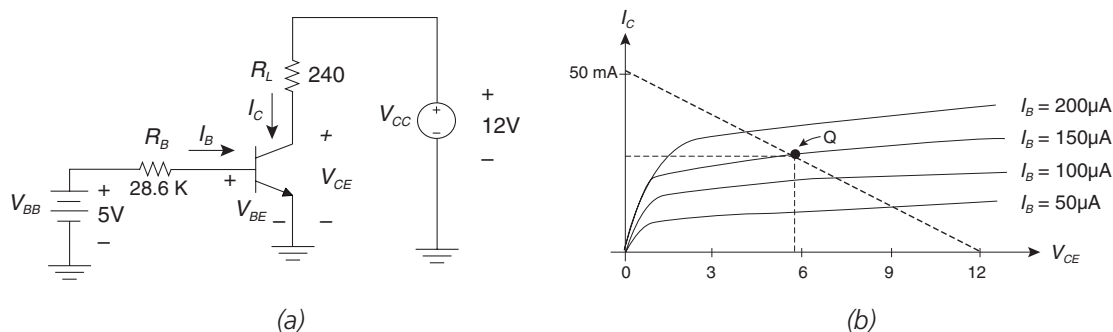


Figure 5-1: Bias design example. (a) Circuit. (b) Load line imposed on the transistor characteristic curves by the 240Ω load resistor. The operating point is labeled Q, and the operating point is $I_C = 26.25 \text{ mA}$, $V_{CE} = 5.7\text{V}$.

This form of simple biasing has several practical design challenges. First, how do you generate the bias supply voltage V_{BB} ? In this example, the main supply $V_{CC} = 12\text{V}$, so a second power supply would be needed to provide the 5V for the base voltage bias.

Secondly, there is the issue of variation in collector current with transistor current gain variations. As shown in the collector current bias current equation, the collector current is proportional to transistor β_F . If β_F varies, the collector current varies proportionally. Shown in **Figure 5-2** are the current-gain curves for a 2N3904 transistor at three different operating temperatures. Therefore, the DC current gain for a transistor varies with temperature and collector current level as shown.

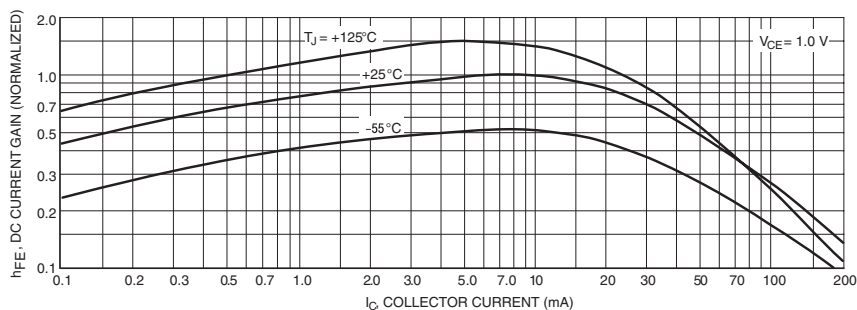


Figure 5-2: Variation in DC current gain β_F (called h_{FE} on this datasheet) for the 2N3904 transistor. (Reprinted with permission of On Semiconductor.)

Third, there is variation in collector current due to transistor V_{BE} variation. For a typical transistor, the base-emitter voltage, measured with a constant collector current bias, decreases approximately two millivolts per degree C of temperature increase.¹

A slightly more complicated biasing arrangement can be implemented to significantly reduce bias variations due to transistor V_{BE} and β_F variations. Consider the circuit of **Figure 5-3a**, where a base bias resistor divider R_{B1} and R_{B2} set the bias voltage at the base of the transistor, and emitter resistor R_E has been added. As we'll see, this combination of components, if properly designed, results in a bias point that is much less dependent on temperature and process variations of β_F and V_{BE} than in the previous circuit.

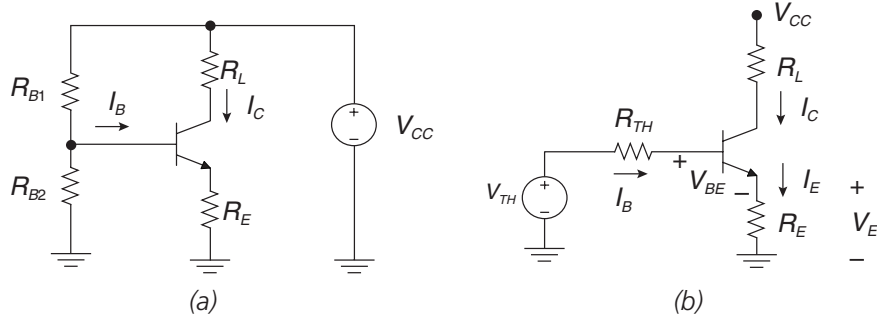


Figure 5-3: Alternative bias circuit. (a) Circuit. (b) Thévenin equivalent circuit.

The Thévenin equivalent of this circuit is shown in **Figure 5-3b** where components are:

$$\begin{aligned} R_{TH} &= R_{B1} \parallel R_{B2} \\ V_{TH} &= V_{CC} \left(\frac{R_{B2}}{R_{B1} + R_{B2}} \right) \end{aligned} \quad [5-3]$$

Using the Thévenin circuit, let's solve for the base current:

$$I_B = \frac{V_{TH} - V_{BE} - V_E}{R_{TH}} \quad [5-4]$$

¹ See, e.g., Robert J. Widlar, "An Exact Expression for the Thermal Variation of the Emitter Base Voltage of Bi-Polar Transistors," *Proceedings of the IEEE*, January 1967, pp. 96–97. Widlar shows that the expression for transistor V_{BE} as a function of temperature is:

$$V_{BE}(T) = V_{Go} \left(1 - \frac{T}{T_o} \right) + V_{BEo} \left(\frac{T}{T_o} \right) + \frac{nkT}{q} \ln \left(\frac{T_o}{T} \right) + \frac{kT}{q} \ln \left(\frac{I_C}{I_{Co}} \right)$$

where V_{Go} = extrapolated bandgap voltage ($\sim 1.2V$ @ $0^\circ K$), T = temperature in $^\circ K$, T_o = reference temperature at V_{BEo} and I_{Co} , and n is a process-dependent constant between 1.5 and 3. At room temperature, the V_{BE} temperature coefficient is approximately -2 mV/ $^\circ K$, or approximately -3000 ppm/ $^\circ C$.

The voltage across the emitter resistor V_E is found by:²

$$V_E = I_E R_E = \left(\frac{\beta_F + 1}{\beta_F} \right) I_C R_E \quad [5-5]$$

Combining these equations results in:

$$I_C = \beta_F \left[\frac{V_{TH} - V_{BE} - \left(\frac{\beta_F + 1}{\beta_F} \right) I_C R_E}{R_{TH}} \right] \approx \beta_F \left[\frac{V_{TH} - V_{BE} - I_C R_E}{R_{TH}} \right] \quad [5-6]$$

Solving for collector current results in:

$$I_C \approx \beta_F \left(\frac{V_{TH} - V_{BE}}{R_{TH}} \right) \left(\frac{1}{1 + \frac{\beta_F R_E}{R_{TH}}} \right) \approx \frac{(V_{TH} - V_{BE})}{\left(R_E + \frac{R_{TH}}{\beta_F} \right)} \quad [5-7]$$

Since we have the extra resistors in this bias arrangement, we have the degrees-of-freedom in our design available to make $V_{TH} \gg V_{BE}$ and $R_E \gg R_{TH}/\beta_F$. Under this set of assumptions, the collector current bias level is approximately:

$$I_C \approx \frac{V_B}{R_E} \quad [5-8]$$

Let's examine intuitively how this circuit provides stabilization of the collector bias current. If collector current I_C increases due to an increase in DC current gain β_F or due to a decrease in V_{BE} , this in turn means that emitter current I_E increases as well. As we note from the base-emitter KVL equation, an increase in I_E results in a decrease in base current. This reduction in base current will in turn stabilize the collector current.

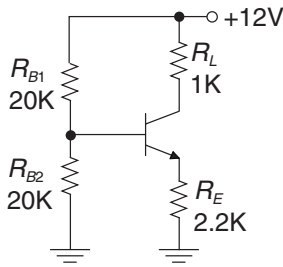
Example 5.1: Biasing example

For the circuit of **Figure 5-4**, let's find the variation in collector current if β_F varies from 40 to 200. Assume that $V_{BE} = 0.7V$.

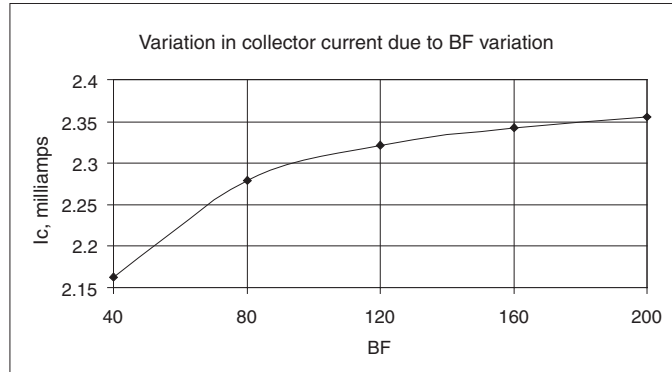
For this circuit, $V_{TH} = 6V$ and $R_{TH} = 10k$. The transistor current gain varies by a factor of $5\times$ as shown, while the collector current shows only a small variation ($\sim 20\%$).

² Note the polarities I have given to the currents, which differ from those found in some texts. We can see that if β_F is large the collector and emitter currents are approximately equal. The actual relationship is:

$$I_B + I_C = I_E \Rightarrow \frac{I_C}{\beta_F} + I_C = I_E \Rightarrow I_E = \left(\frac{\beta_F + 1}{\beta_F} \right) I_C$$



(a)



(b)

Figure 5-4: Bias example. (a) Circuit. (b) Variation in collector current when β_F varies from 40 to 200.

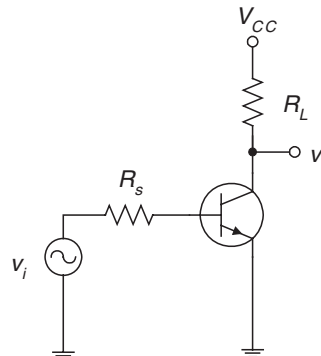
Some Transistor Amplifiers

The common-emitter amplifier

The common-emitter amplifier is a widely used gain block. Using the hybrid pi model, we'll find the bias point, gain and small signal bandwidth of a common-emitter amplifier, built with a 2N3904 transistor (**Figure 5-5**). First we'll do the bandwidth calculations the difficult way, by solving the node equations. For the time being, we'll gloss over the details of the biasing circuit, and assume that the device has the following parameters:

- Transistor $f_T = 300$ MHz; $\omega_T = 1.89 \times 10^9$ radians/second
- Base spreading resistance $r_x = 100\Omega$
- Small-signal current gain $h_{fe} = 150$
- Collector-base junction capacitance $C_\mu = 2$ pF
- Collector current bias $I_C = 2$ mA
- Ambient temperature $T_A = 300$ K

Figure 5-5: Common-emitter amplifier (omitting biasing details). We assume that extra components not shown bias this transistor in the forward-active region. Note that node voltages such as V_{CC} and v_o are measured with respect to ground unless otherwise stated.



The amplifier is driven from a source resistance $R_s = 1 \text{ k}\Omega$ and has a collector resistor $R_L = 1 \text{ k}\Omega$. Our goal in this and the next chapter is to find the gain and the bandwidth by a variety of techniques.

We note that ambient temperature is 300K and hence the thermal voltage $V_T = kT/q \approx 26 \text{ milivolts}$. The transistor is biased on with a collector current of 2 mA, resulting in the following small-signal parameters:

$$\begin{aligned} g_m &= \frac{|I_C|}{V_{TH}} = \frac{2\text{mA}}{26\text{mV}} = 0.077\Omega^{-1} \\ r_\pi &= \frac{h_{fe}}{g_m} = \frac{150}{0.077\Omega^{-1}} = 1950\Omega \\ C_\pi &= \frac{g_m}{\omega_T} = -C_\mu = \frac{0.077}{2\pi(300 \times 10^6)} - 2\text{pF} = 38.8\text{pF} \end{aligned} \quad [5-9]$$

The small-signal model is shown in **Figure 5-6**. The simpler circuit of **Figure 5-6a** will be used to find the low-frequency gain of this amplifier. In order to find the high-frequency bandwidth limit, we'll use the circuit **Figure 5-6b**.

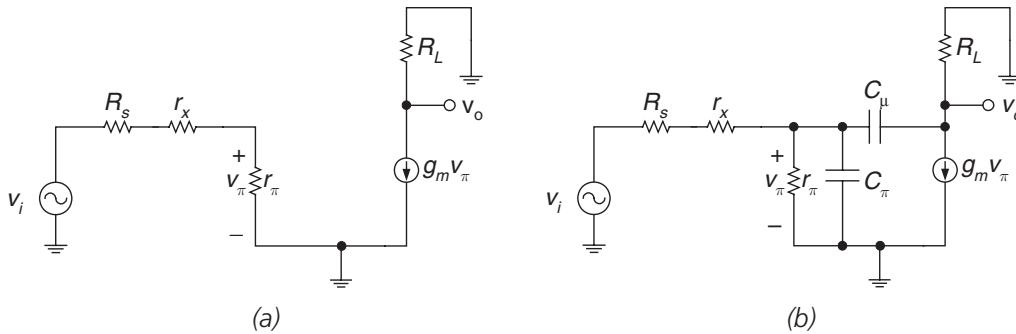


Figure 5-6: Common-emitter amplifier small signal model.
(a) Model valid for low frequencies. (b) Model valid for low and high frequencies.

Low-frequency gain of the common-emitter amplifier

Using the low-frequency small-signal model given in **Figure 5-6a**, let's find the low-frequency gain (at frequencies low enough so that the transistor internal capacitors have no effect). The output voltage is given by:

$$v_o = -g_m v_\pi R_L \quad [5-10]$$

The node voltage v_π is given by the voltage divider:

$$v_\pi = v_i \frac{r_\pi}{R_s + r_x + r_\pi} \quad [5-11]$$

Putting this all together, we can easily find the gain:

$$\frac{v_o}{v_i} = -(g_m R_L) \left(\frac{r_\pi}{R_s + r_x + r_\pi} \right) = -(0.077)(1000) \left(\frac{1950}{1000 + 100 + 1950} \right) \approx -49 \quad [5-12]$$

High-frequency bandwidth limit of common-emitter amplifier

It's considerably more work to find the amplifier response including the effects of transistor internal capacitances C_π and C_μ , shown in **Figure 5-6b**. The node equations at the v_π and v_o nodes are:³

$$\begin{aligned} 1) & (v_i - v_\pi)G'_s - v_\pi(g_\pi + C_\pi s) + (v_o - v_\pi)C_\mu s = 0 \\ 2) & (v_\pi - v_o)C_\mu s - v_o G_L - g_m v_\pi = 0 \end{aligned} \quad [5-13]$$

where:

$$\begin{aligned} G'_s &= \frac{1}{R_s + r_x} \\ g_\pi &= \frac{1}{r_\pi} \\ G_L &= \frac{1}{R_L} \end{aligned} \quad [5-14]$$

Putting this into matrix form results in:

$$\begin{aligned} 1) & -v_\pi [G'_s + g_\pi + (C_\pi + C_\mu)s] + v_o C_\mu s = -v_i G'_s \\ 2) & v_\pi (C_\mu s - g_m) - v_o (G_L + C_\mu s) = 0 \end{aligned} \quad [5-15]$$

or:

$$\begin{bmatrix} -[G'_s + g_\pi + (C_\pi + C_\mu)s] & C_\mu s \\ C_\mu s - g_m & -(G_L + C_\mu s) \end{bmatrix} \begin{bmatrix} v_\pi \\ v_o \end{bmatrix} = \begin{bmatrix} -v_i G'_s \\ 0 \end{bmatrix} \quad [5-16]$$

Using Cramer's rule⁴ to solve for the output voltage is done as follows:

$$v_o = \frac{\det \begin{bmatrix} -[G'_s + g_\pi + (C_\pi + C_\mu)s] & -v_i G'_s \\ C_\mu s - g_m & 0 \end{bmatrix}}{\det \begin{bmatrix} -[G'_s + g_\pi + (C_\pi + C_\mu)s] & C_\mu s \\ C_\mu s - g_m & -(G_L + C_\mu s) \end{bmatrix}} \quad [5-17]$$

³ In many instances the mathematics is a little easier if you use conductances instead of resistance; the conductance of a resistance of value R is given by $G = 1/R$.

⁴ Cramer's rule is discussed in Chapter 16.

where the notation “det” denotes the determinant of the matrix. Solving for the transfer function v_o/v_i results in:

$$\frac{v_o(s)}{v_i(s)} = -(g_m R_L) \left(\frac{G'_s}{G'_s + g_\pi} \right) \left[\frac{1 - \frac{C_\mu}{g_m} s}{\frac{R_L C_\pi C_\mu}{G'_s + g_\pi} s^2 + \frac{1}{G'_s + g_\pi} [R_L (g_m + g_\pi + G'_s) C_\mu + C_\pi + C_\mu] s + 1} \right] \quad [5-18]$$

It's instructive to break this complicated expression down into various terms:⁵

- The $-g_m R_L$ term is what the DC gain would be if the source resistance and base resistance were zero. This is the maximum gain that you can get from a resistively loaded common-emitter amplifier.
- The second gain term is a degradation term due to the effects of r_π loading down the base resistance r_x and source resistance R_s .
- There is a zero in the right-half plane at frequency $+g_m/C_\mu$, which is higher than the ω_T of the transistor. As we'll see later, this zero is at such a high frequency that we can ignore it.
- The denominator shows that there are two poles.

The gain for this amplifier is -49 . MATLAB (**Figure 5-7**) shows that the pole and zero locations are at frequencies:

$$\omega_z = +3.84 \times 10^{10} \text{ radians/second}$$

$$\omega_{p1} = -7.2 \times 10^7 \text{ radians/second}$$

$$\omega_{p2} = -2.5 \times 10^9 \text{ radians/second}$$

Note that both the zero frequency ω_z and the high-frequency pole ω_{p2} are at frequencies higher than the ω_T of the transistor.

⁵ In breaking up the expression into grouped terms, I attempt to show terms grouped in a logical fashion. This follows along with R. D. Middlebrook's concept of “low entropy expressions” (reference at the end of this chapter). We attempt to group terms so that the reader can see the functional dependence of each term in a simple and logical fashion.

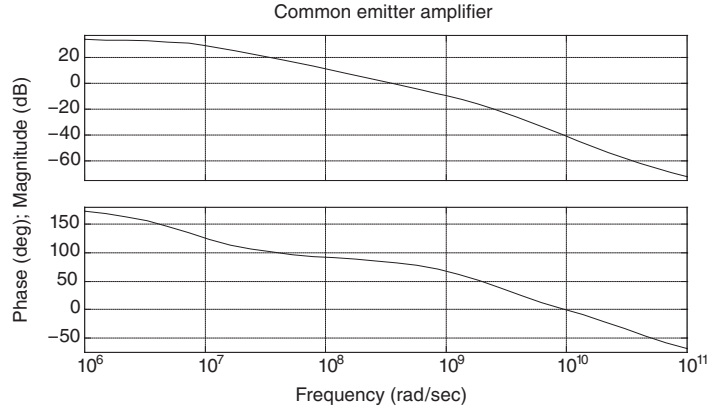


Figure 5-7: Frequency response of common-emitter amplifier.

One-pole approximation for estimating bandwidth of common-emitter amplifier

A fundamental limitation of the hybrid pi model is that it is valid for frequencies significantly lower than the ω_T of the transistor. Given this restriction, the s^2 term in the denominator of the gain expression can be ignored, since:

$$\begin{aligned} |C_\pi C_\mu R_L \omega| &\ll g_m R_L C_\mu \\ \Downarrow \\ \omega &\ll \frac{g_m}{C_\pi} \end{aligned} \quad [5-19]$$

Furthermore, we'll ignore the zero since it is at a frequency higher than ω_T . Since $g_m \gg g_\pi$, we'll ignore g_π wherever possible, resulting in the one-pole approximation:

$$\frac{v_o(s)}{v_i(s)} \approx -g_m R_L \frac{G'_s}{G'_s + g_\pi} \left[\frac{1}{\frac{1}{G'_s + g_\pi} \left[1 + (g_m + G'_s) R_L \right] C_\mu + C_\pi } s + 1 \right] \quad [5-20]$$

Note the insight afforded by this approximation. The denominator term has a single pole, with the C_μ term multiplied by $1 + (g_m + G'_s) R_L$. This term is, in part, equal to the midband gain of the amplifier—hence the multiplication of the feedback capacitance due to the *Miller effect*.⁶ A circuit model illustrating this approximation is shown in **Figure 5-8**. We have lumped the effects of C_π plus the effects of the Miller-effect-modified C_μ into a single equivalent capacitance that we call C_T , with value:⁷

$$C_T = C_\pi + \left[1 + (g_m + G'_s) R_L \right] C_\mu \quad [5-21]$$

⁶ As we'll see later on, the Miller effect comes into play if you have a negative-gain stage with a capacitance across the input-output terminals.

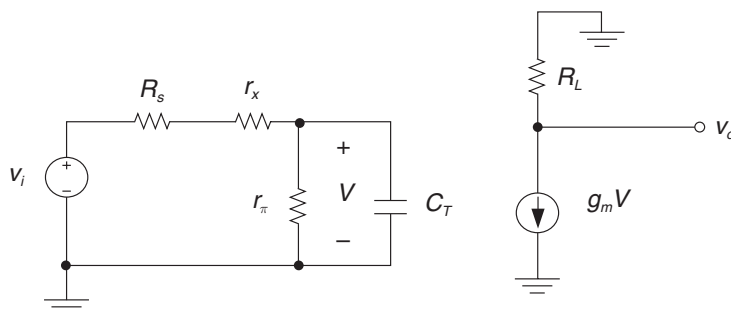


Figure 5-8: Circuit illustrating transistor multiplication caused by the Miller effect.

This approximation predicts a dominant pole at:

$$\omega_{dom} = -\frac{G'_s + g_\pi}{[1 + (g_m + G'_s)R_L]C_\mu + C_\pi} = \frac{G'_s + g_\pi}{C_T} = 7.29 \times 10^6 \text{ rad/sec} \quad [5-22]$$

A plot showing a comparison between the exact result and the Miller approximation is shown in **Figure 5-9**. Note that at frequencies below $\sim 10^9$ radians/second the magnitude responses track very well.

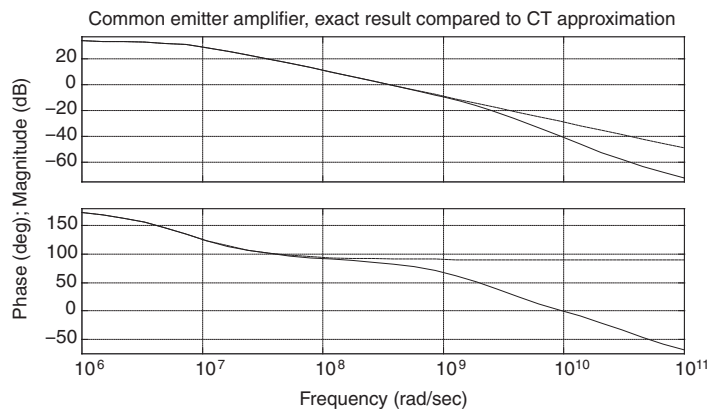


Figure 5-9: Comparison of one-pole (C_T) approximation to exact solution. The exact solution is the solid line; the C_T approximation yields the dashed line.

More discussion of the Miller effect

Using **Figure 5-10**, we can further illustrate the Miller effect.⁸ A capacitor C_f is wrapped around a negative gain of $-A$. By using the circuit of **Figure 5-10b**, we can find the input resistance looking into the input of the amplifier. We have applied a voltage source v_i . The amplifier forces the output voltage to be $-Av_i$.

⁸ See Miller's original 1920 paper, referenced at the end of this chapter.

$$i_t = \frac{v_t - (-Av_t)}{\left(\frac{1}{C_f s} \right)} = v_t (1 + A) C_f s \quad [5-23]$$

The input impedance to the amplifier is the ratio of v_t to i_t , or:

$$Z_{in} = \frac{v_t}{i_t} = \frac{1}{(1 + A) C_f s} \quad [5-24]$$

Therefore, the effect of negative feedback is to make the input capacitance $(1 + A)$ times the feedback capacitance. This effect is known as the *Miller effect* or *Miller multiplication*.

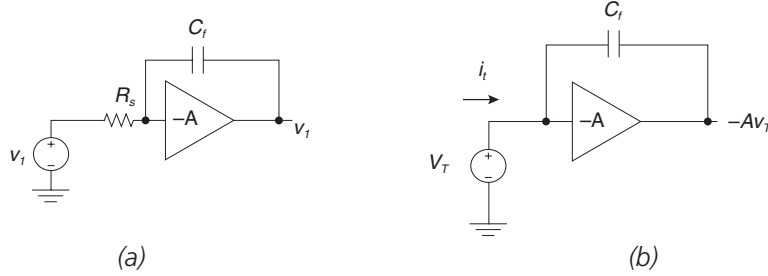


Figure 5-10: Illustration of Miller effect. (a) Circuit. (b) Circuit for determining input capacitance.

Emitter follower gain, input impedance and low-frequency output impedance

The emitter-follower (**Figure 5-11a**) is a buffer stage with high input impedance, low output impedance, and a gain of approximately unity. Using the small-signal low-frequency circuit of **Figure 5-11b**, we'll find the gain, input resistance seen at the base and output resistance seen at the emitter.⁹

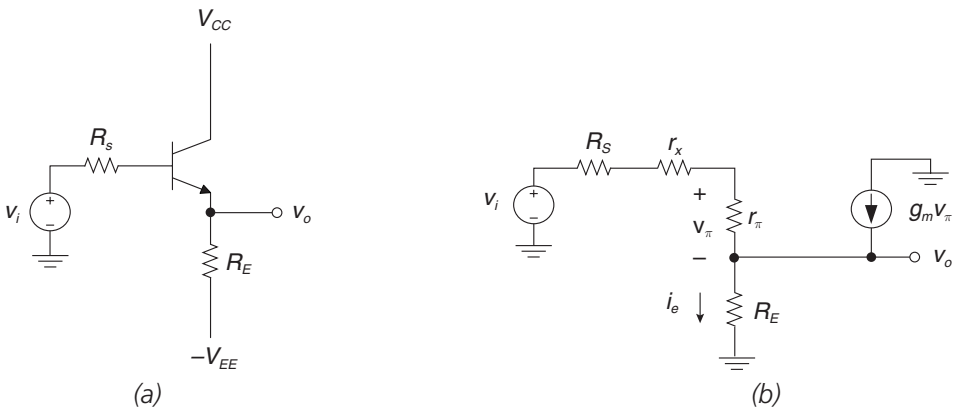
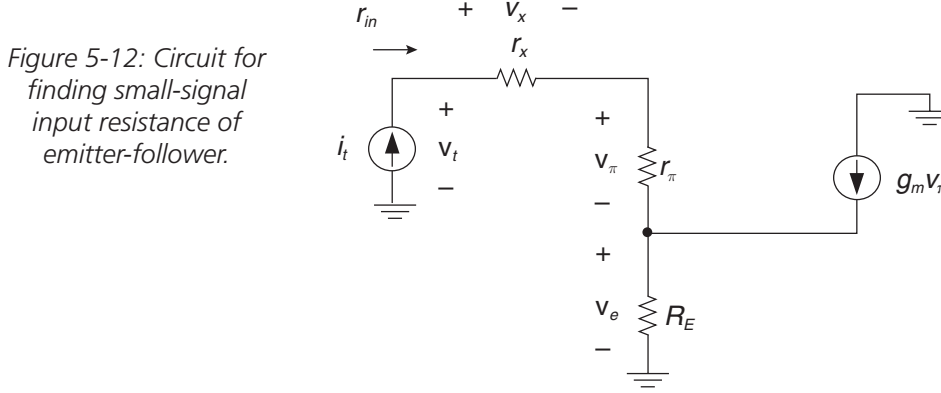


Figure 5-11: Emitter follower. (a) Circuit. (b) Small-signal model.

⁹ Note that the gain, input impedance and output impedance calculations are valid only for low frequencies. In Chapter 7 we'll examine in detail the output impedance of the emitter-follower.

Let's first find the input resistance r_{in} looking into the base of the transistor, using **Figure 5-12**. We have applied a test current source i_t and we'll find the test voltage v_t generated by this current source.



The test voltage v_t is:

$$\begin{aligned}
 v_t &= v_x + v_\pi + v_e = i_t r_x + i_t r_\pi + (i_t + g_m v_\pi) R_E \\
 &\Downarrow \\
 v_t &= i_t r_x + i_t r_\pi + (i_t + g_m (r_\pi i_t)) R_E \\
 &\Downarrow \\
 v_t &= i_t (r_x + r_\pi + (1 + h_{fe}) R_E)
 \end{aligned}
 \tag{5-25}$$

Making note that $g_m r_\pi = h_{fe}$, we find that the input resistance looking into the base of the emitter-follower is:

$$r_{in} = \frac{v_t}{i_t} = r_x + r_\pi + (1 + h_{fe}) R_E
 \tag{5-26}$$

Note that this input resistance can be large for reasonable values of R_E , since R_E is multiplied by h_{fe} in finding the input resistance.

Having found the input resistance looking into the base, we're now at a position to easily find the gain of the emitter-follower. We'll find the gain using the circuit of **Figure 5-11b**. We see that the output voltage that we want to find is given by:

$$v_o = i_e R_E
 \tag{5-27}$$

The incremental emitter current is given by:

$$i_e = \frac{v_\pi}{r_\pi} + g_m v_\pi = (g_\pi + g_m) v_\pi
 \tag{5-28}$$

The voltage v_π is easy to find, since we've previously done the hard work to find the input resistance to the emitter-follower:

$$v_\pi = v_i \left(\frac{r_\pi}{R_s + r_x + r_\pi + (1 + h_{fe})R_E} \right) \quad [5-29]$$

We now find that the gain of the emitter-follower is:

$$\frac{v_o}{v_i} = (g_m + g_\pi) \left(\frac{r_\pi}{R_s + r_x + r_\pi + (1 + h_{fe})R_E} \right) R_E \approx \frac{h_{fe}R_E}{R_s + r_x + r_\pi + (1 + h_{fe})R_E} \quad [5-30]$$

Note that this gain is very close to unity if $h_{fe}R_E \gg R_s, r_x$ and r_π .

Next, we'll find the output resistance r_{out} of the emitter-follower using the circuit of **Figure 5-13**. Initially, we'll ignore the emitter resistance R_E (and put it back in parallel later on). We apply a test voltage source v_t and calculate the resultant test current i_t . The voltage v_π is given by:

$$v_\pi = -v_t \left(\frac{r_\pi}{R_s + r_x + r_\pi} \right) \quad [5-31]$$

The test current i_t is given by:

$$i_t = -g_m v_\pi - \frac{v_\pi}{r_\pi} = -(g_m + g_\pi) v_\pi \quad [5-32]$$

Finally, the test current is:

$$i_t = \left(\frac{g_m r_\pi + 1}{R_s + r_x + r_\pi} \right) v_t \quad [5-33]$$

Therefore, the output resistance of the emitter-follower (remembering to put emitter resistance R_E back into the mix¹⁰) is:

$$r_{out} = \frac{v_t}{i_t} = R_E \left\| \left(\frac{R_s + r_x + r_\pi}{1 + h_{fe}} \right) \right. \quad [5-34]$$

Note that this output resistance is generally small, since the resistors in the base leg are divided by the incremental current gain of the transistor.

¹⁰ In most cases we can ignore R_E altogether, since R_E is usually a large bias resistor, much larger than the resistance looking into the emitter of the transistor.

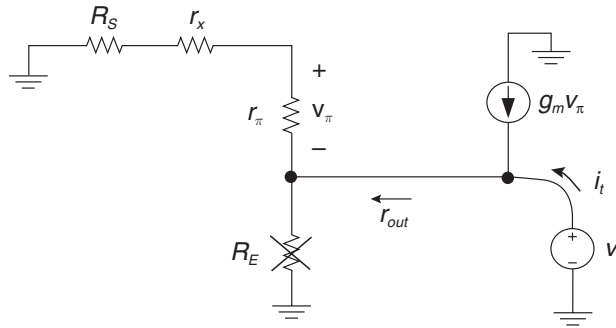


Figure 5-13: Circuit for finding small-signal output resistance of an emitter-follower. We've omitted R_E for the initial calculation. We'll add it back in parallel later on.

Example 5.2: Emitter follower design example

Using the hybrid pi model, we'll find the small-signal gain and bandwidth of an emitter-follower, built with a 2N3904 transistor (**Figure 5-14a**) with the following parameters:

- $f_T = 300 \text{ MHz}$
- $r_x = 100\Omega$
- $h_{fe} = 150$
- $C_\mu = 2 \text{ pF}$

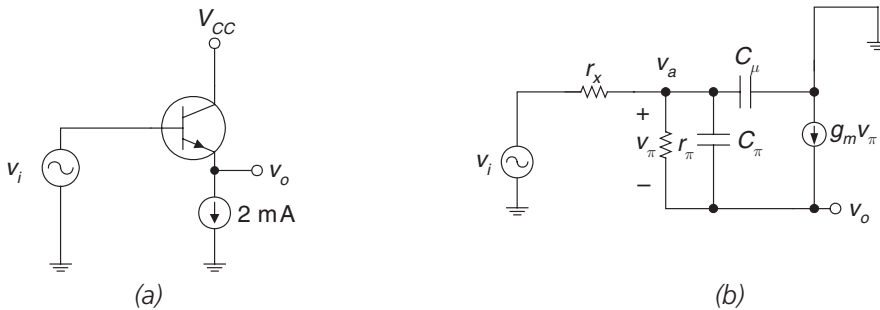


Figure 5-14: Emitter follower design example. (a) Circuit. (b) High-frequency small-signal model.

The emitter-follower is biased with a 2-milliamp current source, resulting in the following small-signal parameters:

$$\begin{aligned}
 g_m &= \frac{|I_C|}{V_{TH}} = \frac{2\text{mA}}{26\text{mV}} = 0.077\Omega^{-1} \\
 r_\pi &= \frac{h_{fe}}{g_m} = \frac{150}{0.077\Omega^{-1}} = 1950\Omega \\
 C_\pi &= \frac{g_m}{\omega_T} - C_\mu = \frac{0.077}{2\pi(300 \times 10^6)} - 2 \text{ pF} = 38.8 \text{ pF}
 \end{aligned}
 \tag{5-35}$$

The small-signal model for the emitter-follower is shown in **Figure 5-14b**. Using this small-signal model, the node equations at the v_a and v_o nodes are:

$$\begin{aligned} 1) & (v_i - v_a)g_x + (v_o - v_a)(g_\pi + C_\pi s) - v_a C_\mu s = 0 \\ 2) & (v_a - v_o)(g_\pi + C_\pi s + g_m) = 0 \end{aligned} \quad [5-36]$$

Putting this into matrix form, and making the approximation that $g_m = h_{fe}g_\pi \gg g_\pi$ results in:

$$\begin{aligned} 1) & -v_a [g_x + g_\pi + (C_\pi + C_\mu)s] + v_o (g_\pi + C_\pi s) = -v_i g_x \\ 2) & v_a (g_m + C_\pi s) - v_o (g_m + C_\pi s) = 0 \end{aligned} \quad [5-37]$$

or:

$$\begin{bmatrix} -[g_x + g_\pi + (C_\pi + C_\mu)s] & g_\pi + C_\pi s \\ g_m + C_\pi s & -(g_m + C_\pi s) \end{bmatrix} \begin{bmatrix} v_a \\ v_o \end{bmatrix} = \begin{bmatrix} -v_i g_x \\ 0 \end{bmatrix}$$

Using Cramer's rule (see Chapter 16) to solve for the transfer function v_o/v_i results in:

$$v_o = \frac{\det \begin{bmatrix} -[g_x + g_\pi + (C_\pi + C_\mu)s] & -v_i g_x \\ g_m + C_\pi s & 0 \end{bmatrix}}{\det \begin{bmatrix} -[g_x + g_\pi + (C_\pi + C_\mu)s] & g_\pi + C_\pi s \\ g_m + C_\pi s & -(g_m + C_\pi s) \end{bmatrix}} \quad [5-38]$$

Solving for v_o/v_i :

$$\frac{v_o(s)}{v_i(s)} \approx \frac{\frac{C_\pi}{g_m}s + 1}{\frac{r_x C_\pi C_\mu}{g_m}s^2 + \left[r_x C_\mu + \frac{C_\pi}{g_m} \right]s + 1} \quad [5-39]$$

Note that the denominator is exactly in the form:

$$(\tau_1 s + 1)(\tau_2 s + 1) = \tau_1 \tau_2 s^2 + (\tau_1 + \tau_2)s + 1 \quad [5-40]$$

Therefore, the two poles are at frequencies:

$$\begin{aligned} \omega_{p1} &= -\frac{1}{\tau_1} = -\frac{g_m}{C_\pi} = -1.95 \times 10^9 \text{ rad/sec} \\ \omega_{p2} &= -\frac{1}{\tau_2} = -\frac{1}{r_x C_\mu} = -5 \times 10^9 \text{ rad/sec} \end{aligned} \quad [5-41]$$

Note that there is also a zero at C_π/g_m , which in this simplified model exactly cancels the pole at $-\omega_{p1}$. Therefore, the bandwidth for this circuit is:

$$\omega_h \approx -\omega_{p2} = 5 \times 10^9 \text{ rad/sec}$$

$$f_h \approx \frac{\omega_h}{2\pi} = 795 \text{ MHz} \quad [5-42]$$

which matches closely the SPICE result. Actually, the pole and zero don't exactly cancel. This is an artifact of our ignoring g_π as compared to g_m . Also, the preceding result must be taken with a grain of salt, as the model predicts a bandwidth in excess of the f_T of the transistor.

A SPICE result showing the gain and bandwidth is shown in **Figure 5-15**. This result shows that the gain = 1 (as expected) and the -3dB bandwidth is approximately 800 MHz.

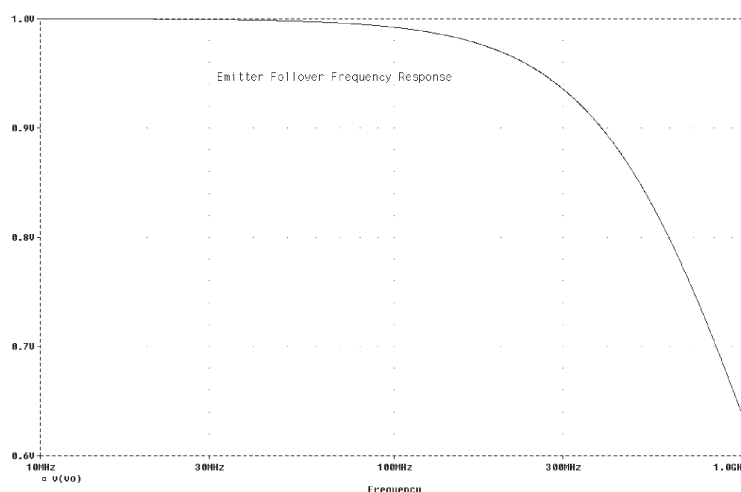


Figure 5-15: SPICE result for emitter-follower, showing bandwidth of approximately 800 MHz.

Differential amplifier

The transistor differential amplifier (**Figure 5-16**) is a ubiquitous building block used commonly as the front-end of an operational amplifier. The differential amplifier is used to amplify the difference between its two inputs, while rejecting the DC value common to the two inputs.¹¹

If we do KVL around the voltage source and base-emitter junctions, we find:

$$-V_1 + V_{BE1} - V_{BE2} + V_2 = 0 \quad [5-43]$$

¹¹ Said another way, the differential amplifier is designed to have high “differential-mode” gain and very low “common-mode” gain.

From the ideal transistor relationships and with $V_{BE} \gg kT/q$ (~26 millivolts at room temperature) we find:

$$\begin{aligned} I_{C1} &\approx I_s e^{\frac{qV_{BE1}}{kT}} \Rightarrow V_{BE1} \approx \frac{kT}{q} \ln \left(\frac{I_{C1}}{I_s} \right) \\ I_{C2} &\approx I_s e^{\frac{qV_{BE2}}{kT}} \Rightarrow V_{BE2} \approx \frac{kT}{q} \ln \left(\frac{I_{C2}}{I_s} \right) \end{aligned} \quad [5-44]$$

Combining the equations above results in an equation relating the ratio of the two collector currents:

$$\frac{I_{C1}}{I_{C2}} = e^{\frac{q(V_1 - V_2)}{kT}} = e^{\frac{qV_{id}}{kT}} \quad [5-45]$$

We have defined the differential input voltage as the difference between the two inputs, or $V_{id} = V_1 - V_2$. Now, if $\beta_F \gg 1$ then $I_E \approx I_C$ for each transistor and $I_{C1} + I_{C2} \approx I_{BIAS}$, and we can write:

$$\begin{aligned} I_{C2} &\approx \frac{I_{BIAS}}{1 + e^{\frac{qV_{id}}{kT}}} \\ I_{C1} &\approx \frac{I_{BIAS}}{1 + e^{-\frac{qV_{id}}{kT}}} \end{aligned} \quad [5-46]$$

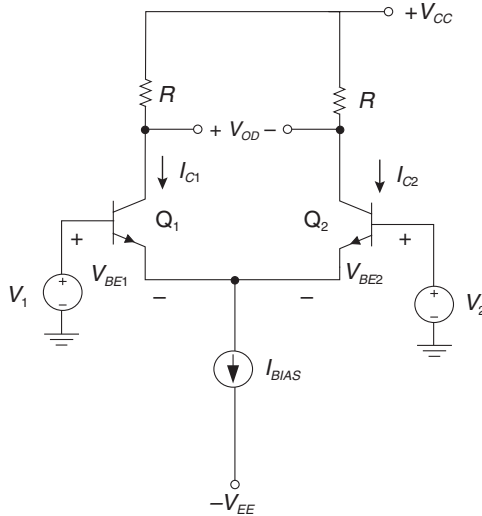


Figure 5-16: Ideal transistor differential amplifier. Note that the differential input voltage V_{id} is the difference between V_1 and V_2 , or $V_{id} = V_1 - V_2$.

The exponential terms mean that the range of V_{id} over which I_{C1} and I_{C2} vary is only a few kT/q . We can also find the differential output voltage V_{od} by recognizing that V_{od} is the difference between the two transistor collector voltages. In **Figure 5-17b**, we implicitly assume that the load resistors R are small enough so that Q_1 and Q_2 do not saturate. The range of voltages over which approximately linear amplification occurs is on order of $|V_{id}| < 25$ millivolts. Note that the differential output voltage depends only on the differential input voltage V_{id} , and *not* on other design parameters such as bias current level I_{BIAS} or power supply voltage V_{CC} . This is one of the advantages of the fully differential amplifier topology.

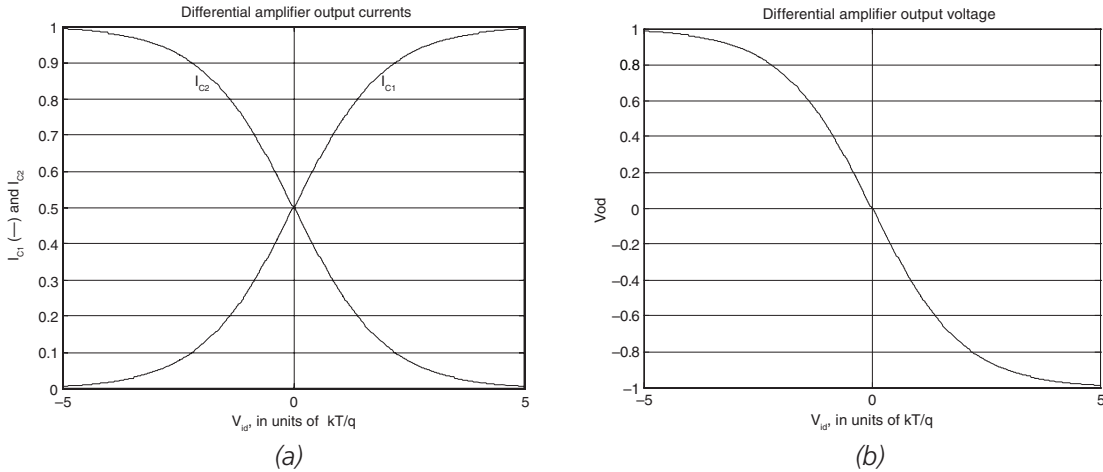


Figure 5-17: Ideal transistor differential amplifier outputs.

(a) Output currents I_{C1} and I_{C2} in units of I_{BIAS} variation as V_{id} varies from $-5 kT/q$ to $+5 kT/q$. (b) Differential output voltage V_{od} in units of $I_{BIAS}R$.

To analyze the differential gain, we make use of a *half-circuit* technique. Let's assume that we're operating in a mode when both Q_1 and Q_2 are biased ON, and hence $V_1 - V_2$ is a small voltage less than a few kT/q . In this mode, the differential input voltage is small enough so that both transistors carry approximately the same current (or $\sim I_{BIAS}/2$). In the differential mode of operation, when V_1 moves up, V_2 moves down, and vice versa. Therefore, the emitters of the coupled pair remain at incremental ground during this operation. Therefore, we can analyze the left side and the right side of the circuits independently. A circuit valid for the differential mode is shown in **Figure 5-18a**.¹² Since the voltage at the symmetry plane doesn't move when we're in the differential mode, we can ground the symmetry plane, resulting in

¹² Note that in order to further illustrate the symmetry of the circuit, we've broken the bias generator into two equal current sources of value $I_{BIAS}/2$.

the incremental circuit of **Figure 5-18b**¹³ valid for differential-mode operation. From this circuit, we find that the differential-mode incremental gain ($A_{V,DM}$) is:

$$\begin{aligned} \frac{v_{od}}{v_{id}} &= A_{V,DM} = -g_m R \\ g_m &= \frac{I_{BIAS}}{2V_T} = \frac{I_{BIAS}}{2 \left(\frac{kT}{q} \right)} \end{aligned} \quad [5-47]$$

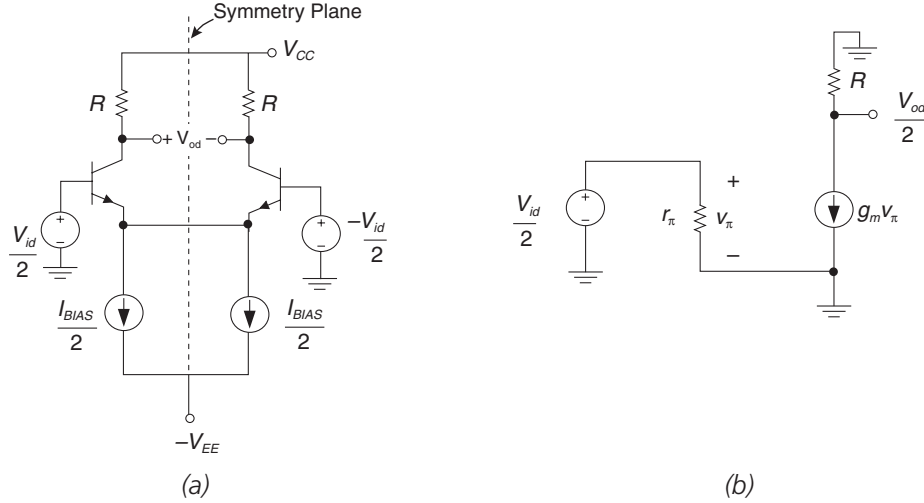


Figure 5-18: Illustration of differential mode operation. (a) Differential mode operation showing symmetry plane. (b) Small-signal model of differential-mode half-circuit.

A real-world differential amplifier has an important parasitic element—the finite incremental output resistance of the current source. This is shown in **Figure 5-19a**, where we see that the I_{BIAS} current source has a finite output resistance R_{CS} . We'll see that this finite resistance affects how well the amplifier rejects *common-mode* signals. Common-mode signals are signals that are common to both inputs, as shown in **Figure 5-19b**. The same input v_{ic} is applied to the two transistor inputs. Note that in this figure we have redrawn the circuit showing the symmetry plane.

Again, we can exploit symmetry in order to simplify circuit analysis. Note that the same input is applied to the bases of both transistors and that the circuit is symmetric. Therefore, there is no current across the symmetry plane (i.e., current $i_x = 0$). Therefore, we're free to cut the circuit apart at the symmetry plane. This results in the small-signal model of **Figure 5-19c**.

¹³ For simplicity, we assume that base spreading resistance $r_x = 0$.

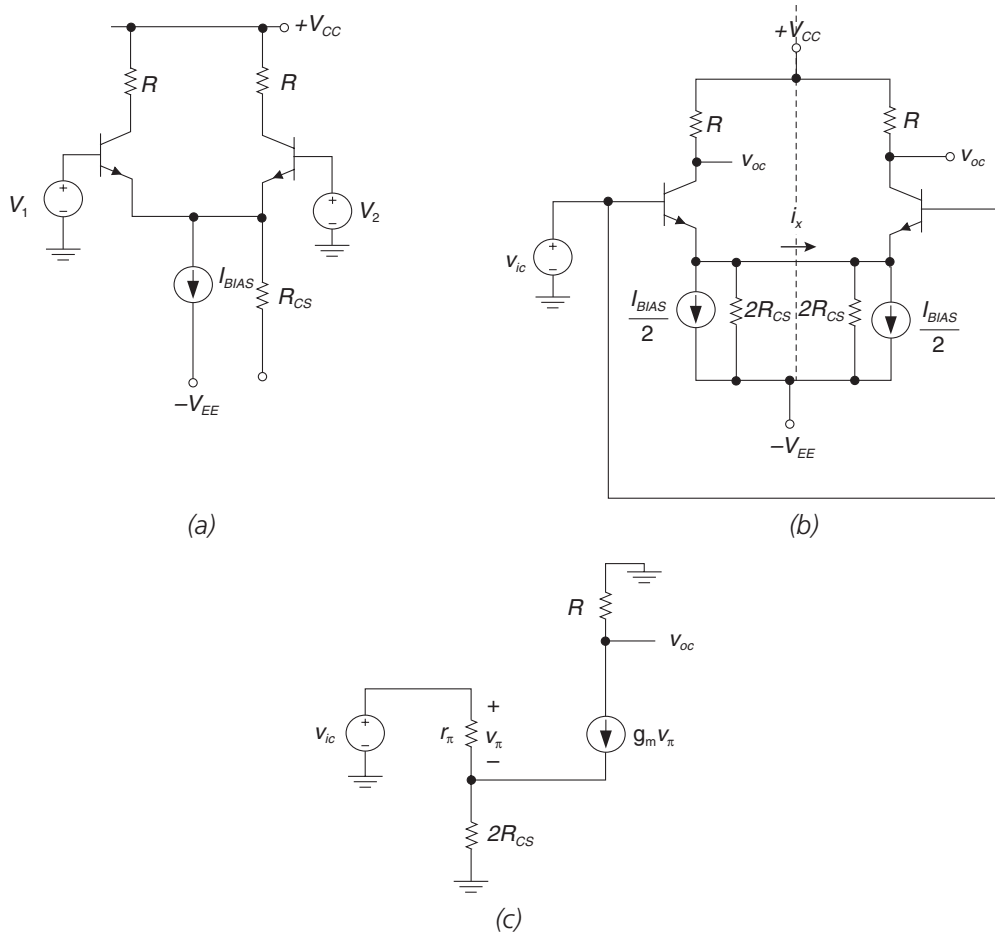


Figure 5-19: Differential amplifier showing a parasitic element that affects common-mode gain. (a) Bias current source with finite output resistance R_{CS} . (b) Circuit redrawn showing common-mode symmetry plane. (c) Circuit valid for common mode analysis.

Using **Figure 5-19c**, we can find the common-mode gain $A_{V,CM}$, skipping over the mathematical manipulations:¹⁴

$$\begin{aligned} \frac{v_{oc}}{v_{ic}} &= A_{V,CM} = -(g_m R) \left(\frac{1}{1 + 2g_m R_{CS}} \right) \\ g_m &= \frac{I_{BIAS}}{2V_T} \end{aligned} \quad [5-48]$$

An important figure of merit is the ratio of differential-mode gain to common-mode gain. In a differential amplifier, we want a high differential mode gain and a very low (ideally zero) common-mode gain. The common mode rejection ratio (CMRR) of a differential amplifier is given by:

$$CMRR = \frac{|A_{V,DM}|}{|A_{V,CM}|} = 1 + 2g_m R_{CS} \quad [5-49]$$

Therefore, to achieve high CMRR we need to bias the differential pair with a current source with high output resistance.

An analogy to help explain differential and common modes of operation

A simple mechanical analogy (**Figure 5-20**) can illustrate differential and common mode operation.¹⁵ Let's assume that we have a teeter-totter, with the motion of the ends of the teeter-totter being analogous to the voltage drive levels at the inputs of the differential amplifier. In the differential mode of operation (**Figure 5-20a**), the center of the teeter-totter is fixed, and one input goes UP and the other goes DOWN the same amount. This motion of the ends of the teeter-totter is analogous to the variation in the inputs to the differential amplifier in the differential mode.

¹⁴ We can use the previous work we did on the input resistor of the emitter-follower to arrive at this result. The input resistance at the base of the transistor is:

$$r_\pi + (1 + h_{fe})(2R_{CS})$$

We can now find v_π :

$$v_\pi = \frac{r_\pi}{r_\pi + (1 + h_{fe})(2R_{CS})}$$

Next, we find the common-mode gain:

$$\frac{v_{oc}}{v_{ic}} = \frac{-g_m R r_\pi}{r_\pi + (1 + h_{fe})(2R_{CS})} = \frac{-g_m R}{1 + \frac{(1 + h_{fe})(2R_{CS})}{r_\pi}} \approx \frac{-g_m R}{1 + 2g_m R_{CS}}$$

¹⁵ Thanks go to Prof. John McNeill at Worcester Polytechnic Institute, whose course notes on undergraduate IC design use this analogy.

In the common mode of operation (**Figure 5-20b**), the center of the teeter-totter is allowed to move up and down, but both ends go UP or DOWN equally. This is analogous to the common mode of operation when a common voltage drives the differential amplifier.

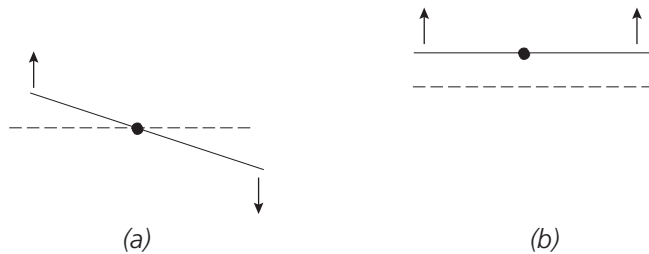


Figure 5-20: Teeter-totter analogy to differential and common mode operation.
(a) Differential mode. (b) Common mode.

Example 5.3: Peaking amplifier

Shown in **Figure 5-21a** is an inductively peaked amplifier. Assume that the transistor has parameters: $\omega_T = 2\pi \times 800 \times 10^6$ radians/sec; $h_{fe} = 100$; $C_\mu = 0.7$ pF, $r_x \sim 0$.¹⁶ As we have seen earlier, the gain of the common-emitter amplifier rolls-off at a finite frequency due to the bandwidth limiting effects of C_π and C_μ . One way to counteract this bandwidth limit is to use an inductor in series with the load resistor. At high frequencies, the impedance of the load increases. Using this technique the bandwidth can be extended.

The small-signal model of this amplifier is shown in **Figure 5-21b**, assuming operation at a frequency high enough where the emitter bypass capacitor behaves as a short-circuit. If we assume that r_x is very small, we can further simplify the small-signal circuit as in **Figure 5-21c**. The small-signal parameters are:

$$\begin{aligned}
 g_m &= \frac{|I_C|}{V_{TH}} = \frac{5 \text{ mA}}{26 \text{ mV}} \approx 0.19 \text{ A/V} \\
 r_\pi &= \frac{h_{fe}}{g_m} = \frac{100}{0.19} = 526 \Omega \\
 C_\pi &= \frac{g_m}{\omega_T} - C_\mu = \left(\frac{0.19}{2\pi \times 800 \times 10^6} \right) - 0.7 \text{ pF} = 37 \text{ pF}
 \end{aligned} \tag{5-50}$$

Using the small-signal model, we can find the midband gain as:

$$A_v = -g_m R_L = -190 \tag{5-51}$$

¹⁶ Ignoring r_x leads to a mathematical result that is relatively easy to understand. As we'll see in a future chapter (Chapter 7), r_x can often be a dominant bandwidth limitation in circuits that have a low source resistance. So, ignore r_x at your peril!

Using the small-signal model of **Figure 5-21c**, we can find the overall gain of this amplifier by writing node equations. Let's sum currents at node v_o :

$$(v_i - v_o)C_\mu s - g_m v_i - \frac{v_o}{R_L + L_{pk}s} = 0 \quad [5-52]$$

Solving for the input-output transfer function results in:

$$\frac{v_o}{v_i} = -g_m R_L \frac{\left(1 + \frac{L_{pk}}{R_L} s\right) \left(1 - \frac{C_\mu}{g_m} s\right)}{L_{pk} C_\mu s^2 + R_L C_\mu s + 1} \approx -g_m R_L \frac{\left(1 + \frac{L_{pk}}{R_L} s\right)}{L_{pk} C_\mu s^2 + R_L C_\mu s + 1} \quad [5-53]$$

For $L_{pk} = 0$, this transfer function predicts a bandwidth of $1/(R_L C_\mu)$, or $2\pi \times 227 \times 10^6$ radians/second. This is confirmed by PSPICE (**Figure 5-22**), which also shows that the midband voltage gain is -190 . With an inductor $L_{pk} = 0.25 \mu\text{H}$, the -3dB bandwidth extends to approximately 390 MHz, with a little bit of gain peaking.¹⁷ With larger values of L_{pk} the response becomes very peaky, as shown in the plots.

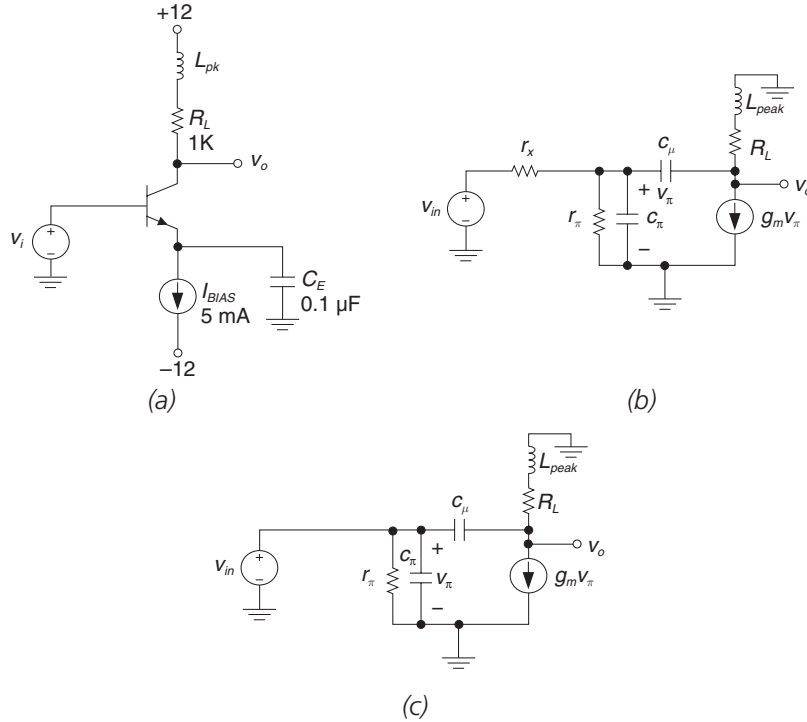


Figure 5-21: Inductively peaked amplifier. (a) Circuit. (b) Small-signal model.
(c) Simplified small-signal model assuming that r_π is very small.

¹⁷ Detailed mathematical analyses of similar circuits is given in Tom Lee's *The Design of CMOS Radio-Frequency Circuits*, pp. 178–184.

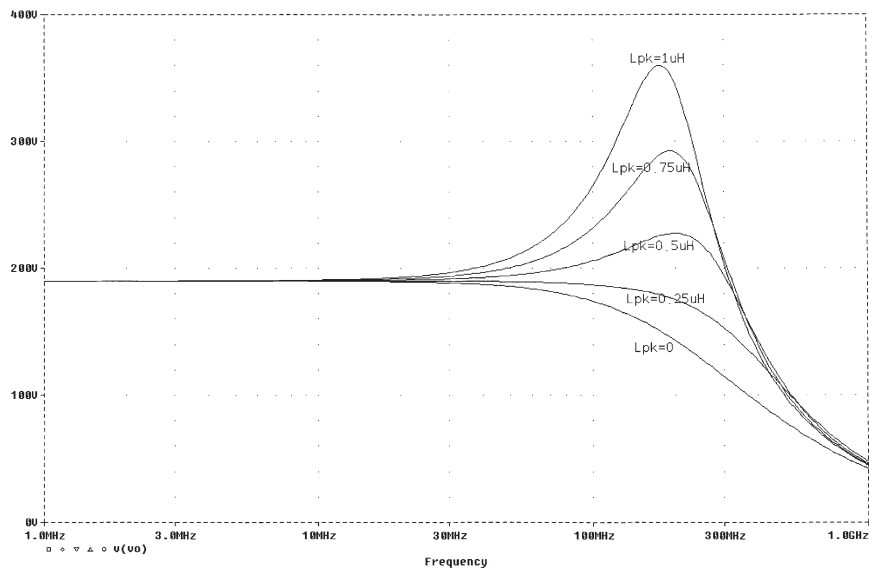


Figure 5-22: SPICE analysis of inductively peaked amplifier.
 Plot of gain, for $L_{pk} = 0, 0.25 \mu H, 0.5 \mu H, \dots 1.0 \mu H$

Chapter 5 Problems

Problem 5.1

For the circuit of **Figure 5-4**, find the variation in collector current if V_{BE} varies from 0.6V to 0.8V. Assume that current gain remains constant at $\beta_F = 100$.

Problem 5.2

For the circuit in **Figure 5-23**, calculate bias point values:

- (a) I_{C1}
- (b) I_{C2}
- (c) V_{E1}
- (d) V_{E2}

Throughout, assume that $V_{BE} = 0.7\text{V}$ and $\beta_F = 100$.

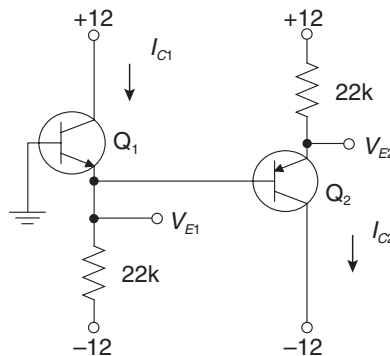
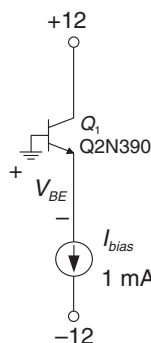


Figure 5-23: Transistor circuit for Problem 5.2.

Problem 5.3

The emitter-follower circuit of **Figure 5-24** is biased at a constant emitter current of 1 milli-amp. Using PSPICE, find the variation in transistor V_{BE} as the ambient temperature rises from 25°C to 75°C.

Figure 5-24: Circuit for Problem 5.3.

**Problem 5.4**

Using reasonable engineering approximations, calculate the bias point values I_C and V_{CE} for the transistor amplifier in **Figure 5-25**.

Problem 5.5

- Using the incremental model calculate the “midband” gain of the transistor amplifier in **Figure 5-25**, v_o/v_{in} . Midband is the frequency range high enough so that C_c and C_E behave as short circuits, but low enough so that the transistor internal capacitances have little or no effect.
- Using the one-pole approximation (Miller approximation), calculate the high-frequency -3dB bandwidth f_H (in Hz).
- Sketch the Bode plot of the magnitude response of v_o/v_{in} , but do not calculate the low frequency breakpoint which is dependent on coupling and bypass capacitors C_c and C_E (more on this later). Plot f_h in Hz.
- Simulate your circuit using SPICE, and compare your results to calculations. In your SPICE calculation, don’t use the SPICE 2N3904 model; instead, directly input the hybrid pi model you derive above.

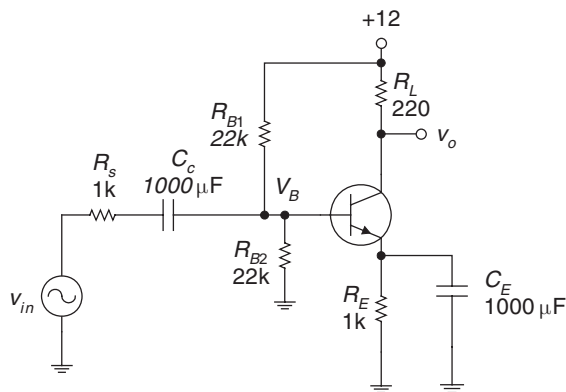


Figure 5-25: Common emitter amplifier with biasing network.

Problem 5.6

For the transistor circuit of **Figure 5-26**, calculate:

- The operating point current I_c and the DC value of V_{out} .
- Small signal parameters r_π and g_m .
- Draw the mid-frequency small-signal model, assuming C_{in} is a short circuit. Assume that the frequency is low enough so that the transistor internal capacitances have no effect.
- Calculate the AC input resistance looking into the base terminal of the capacitor (assuming C_{in} acts as a short circuit at signal frequencies).
- Find AC voltage gain $A_v = |v_{out}/v_{in}|$ (again, assuming that C_{in} acts as a short circuit). (*Hint: there's a clever way to use the results of part (d) to greatly simplify this result.*)
- Calculate the low-frequency breakpoint of A_v . (Hint: at DC, the gain of this circuit is zero, due to C_{in} . At high frequency, the gain is what you calculate in part (e). The breakpoint is where the gain has risen to -3dB from maximum value). Use the result from part (c) to find this.
- Sketch (but don't calculate a high-frequency breakpoint) a Bode plot (magnitude only) of the gain.

Throughout the problem, make reasonable approximations and justify them.

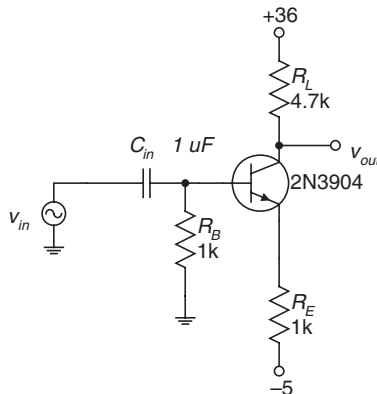


Figure 5-26: Circuit for Problem 5.6.

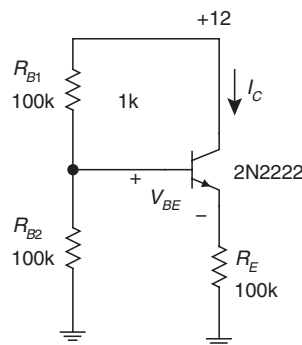
Problem 5.7

For the transistor circuit of **Figure 5-27**,

- Estimate collector current at -55°C , 25°C and 125°C , using information from the 2N2222 datasheet as a guide, and a closed-form solution for collector current. Assume that $V_{BE} = 0.7\text{V}$ at 25°C , and that the transistor V_{BE} has a temperature coefficient of $-2\text{ mV}/^\circ\text{C}$.

- (b) Simulate using PSPICE at the three different ambient temperatures. Comment on how well the PSPICE result matches up with what you calculated in part (a). Also, comment on the bias stability of this circuit and how you might improve the bias stability.

Figure 5-27: Circuit for Problem 5.7.

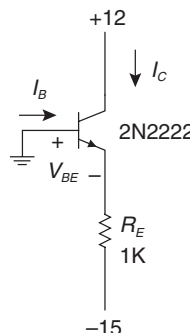


Problem 5.8

For the circuit of **Figure 5-28**, assume that $r_x = 20\Omega$ and that $V_{BE} = 0.7$ volts.

- Using the 2N2222 datasheet, find reasonable values for DC current gain h_{FE} and small-signal current gain h_{fe} .
- Find collector current I_C and base current I_B .
- Find small-signal parameters g_m and r_π .
- Draw the low-frequency incremental circuit.

Figure 5-28: NPN transistor circuit built with 2N2222.



Problem 5.9

In the 2N3904 circuit in **Figure 5-29**, the collector current will vary with temperature due to the dependencies of V_{BE} and β_F calculated above.

- Calculate the operating point values of collector current I_C and base-emitter voltage V_{BE} at 25°C, 50°C and 75°C. You may assume that $V_{BE} = 725$ mV at $T_A = 25^\circ\text{C}$. Assume that the temperature dependence of V_{BE} in a transistor at constant current is approximately -2.2 mV per degree C, and the temperature dependence of β_F is in the neighborhood of $+7000$ PPM/C.

- (b) Now, simulate using SPICE and compare your values to that calculated in part (a) above. Can you draw any assumptions about the relative bias stability of this circuit? How might you improve the bias stability?

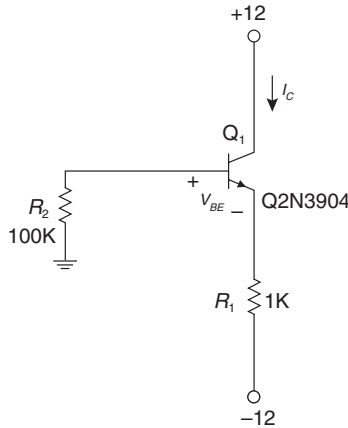


Figure 5-29: Transistor circuit for Problem 5.9.

References

- Chuang, C. T., "Analysis of the settling behavior of an operational amplifier," *IEEE Journal of Solid State Circuits*, vol. 17, no. 1, Feb. 1982, pp. 74–80.
- Filipkowski, A., "Poles and zeros in transistor amplifiers introduced by Miller effect," *IEEE Transactions on Education*, vol. 42, no. 4, Nov. 1999, pp. 349–351.
- Gilbert, Barrie, "All You Ever Need to Know About Bandgaps," lecture, November 16, 1978.
- Gray, Paul E., and Searle, Campbell L., *Electronic Principles Physics, Models and Circuits*, John Wiley, 1969.
- Gray, Paul R., and Meyer, Robert G., *Analysis and Design of Analog Integrated Circuits*, 2d edition, John Wiley, 1984.
- Knapp, Ron, "Selection criteria assist in choice of optimum reference," *EDN*, February 18, 1988, pp. 183–192.
- , "Back-to-basics approach yields stable references," *EDN*, June 9, 1988, pp. 193–198. *Detailed description of bandgap and Zener voltage references.*
- Lee, Thomas H., *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge University Press, 1998.
- Lindmayer, J., and North, W., "The Inductive Effect in Transistors," *Solid-State Electronics*, vol. 8, 1965, pp. 409–415.
- Mercer, M. J., and Burns, S. G., "High-frequency broadband amplifier ASIC design optimization using pole-zero compensation techniques," *IEEE International Symposium on Circuits and Systems*, May 1–3, 1990, pp. 3225–3229.

- Middlebrook, R. D., “Low-entropy expressions: the key to design-oriented analysis,” *Proceedings of the Twenty-First Annual Conference “Engineering Education in a New World Order,”* September 21–24, 1991, pp. 399–403.
- Miller, John M., “Dependence of the input impedance of a three-electrode vacuum tube upon the load on the plate circuit,” *Scientific Papers of the Bureau of Standards*, vol. 15, no. 351, pp. 367–385, 1920.
- Muller, Richard S., and Kamins, Theodore I., *Device Electronics for Integrated Circuits*, 2d edition, John Wiley, 1986.
- Neudeck, Gerold W., *Modular Series on Solid State Devices, The Bipolar Junction Transistor*, Addison-Wesley, 1983.
- Pease, Bob, “The Design of Band-Gap Reference Circuits: Trials and Tribulations,” *IEEE Proceedings of the 1990 Bipolar Circuits and Technology Meeting*, September 17–18, 1990, Minneapolis, Minnesota.
- Rincón-Mora, Gabriel Alfonso, *Voltage References*, IEEE, 2002.
- Searle, C. L., Boothroyd, A. R., Angelo, E. J., Jr., Gray, P. E., and Peterson, D. O., *Elementary Circuit Properties of Transistors*, (SEEC Volume 3), John Wiley, 1964.
- Sze, S. M., *Physics of Semiconductor Devices*, 2d edition, John Wiley, 1981.
- Thompson, Marc T., “Design Linear Circuits Using OCTC Calculations,” *Electronic Design (Special Analog Issue)* June 24, 1993, pp. 41–47.
- , “SCTC Analysis Estimates Low-Frequency –3-dB Point,” *Electronic Design*, October 1, 1993, pp. 65–68.
- , “Network Tricks Aid in OCTC,” *Electronic Design*, December 16, 1993, pp. 67–70.
- , “Tips for Designing High-Gain Amplifiers,” *Electronic Design*, May 16, 1994, pp. 83–90.
- Thornton, R. D., DeWitt, D., Chenette, E. R., and Gray, P. E., *Characteristics and Limitations of Transistors*, (SEEC Vol. 4), John Wiley, 1966.
- Thornton, R. D., Searle, C. L., Pederson, D. O., Adler, R. B., and Angelo, E. J., Jr., *Multistage Transistor Circuits* (SEEC Vol. 5), John Wiley, 1965.
- Widlar, Robert J., “New Developments in IC Voltage Regulators,” *IEEE Journal of Solid-State Circuits*, vol. SC-6, no. 1, February 1971, pp. 2–7.
- , “An Exact Expression for the Thermal Variation of the Emitter Base Voltage of Bipolar Transistors,” *Proceedings of the IEEE*, January 1967, pp. 96–97. (Widlar was the inventor of the National Semiconductor LM113 reference diode, which is a monolithic band-gap voltage reference introduced in 1971.)
- Wing-Hung, Ki, Der, L., and Lam, S., “Re-examination of pole splitting of a generic single stage amplifier,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 44, no. 1, Jan 1997.
- Yahya, C. B., “Design of wideband low noise transimpedance amplifiers for optical communications,” *Proceedings of the 43rd IEEE Midwest Symposium on Circuits and Systems*, August 8–11, 2000, pp. 804–80.

Bandwidth Estimation Techniques and the Method of Open-Circuit Time Constants

In This Chapter

- It's possible to use PSPICE or other circuit simulators to determine the gain and bandwidth of transistor amplifiers, or for that matter any circuit with lumped elements and dependent sources. However, while PSPICE is a useful tool for analyzing circuits, it doesn't give that much insight and intuition into amplifier design. This chapter covers "back of the envelope" techniques for estimating the bandwidth of transistor amplifiers. The techniques in this chapter are employed on bipolar transistor amplifiers, but can also be used for CMOS.

Introduction to Open-Circuit Time Constants

The method of open-circuit time constants (OCTC) is a powerful approximate analysis tool that allows estimation of the high-frequency -3dB bandwidth of circuits containing resistors, capacitors and dependent sources. The method was developed by R. B. Adler and others at MIT.¹ and is a valuable design tool in finding the root cause of bandwidth limitation in circuits. The usefulness of the technique lies in the fact that it allows identification of local bandwidth limitations in amplifiers.

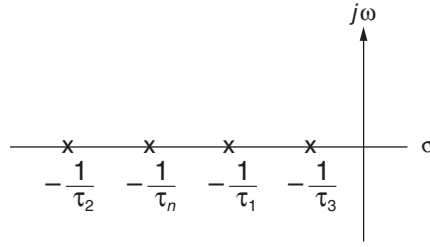
Following a brief mathematical discussion of the method of open-circuit time constants, the method is applied to several different amplifier topologies. The open-circuit time constant results are compared to closed-form solutions and SPICE simulations.

At first, let's consider a system with n real-axis poles (**Figure 6-1**). The transfer function for this system is:

$$H(s) = \frac{1}{(\tau_1 s + 1)(\tau_2 s + 1) \cdots (\tau_n s + 1)} \quad [6-1]$$

¹ See, e.g., the Semiconductor Electronics Education Committee's *Multistage Transistor Circuits*, volume 5, by Thornton, Searle, Pederson, Adler and Angelo.

Figure 6-1: System with n negative real-axis poles.



Expanding the denominator results in:

$$H(s) = \frac{1}{(\tau_1 \tau_2 \cdots \tau_n) s^n + \cdots + (\tau_1 + \tau_2 + \cdots + \tau_n) s + 1} \quad [6-2]$$

We note that the first-order s term has as its coefficient the sum of reciprocal pole locations. In a simpler case, let's look at a system with two negative real-axis poles, with transfer function:

$$H(s) = \frac{1}{\tau_1 \tau_2 s^2 + (\tau_1 + \tau_2) s + 1} \quad [6-3]$$

In using the method of open-circuit time constants, we throw away the higher-order s terms. The method of open-circuit time constants approximates the second-order transfer function as:

$$H(s) \approx \frac{1}{(\tau_1 + \tau_2) s + 1} \quad [6-4]$$

Using the method of open-circuit time constants, our estimate of bandwidth is:

$$\omega_h \approx \frac{1}{(\tau_1 + \tau_2)} \quad [6-5]$$

When is this approximation justified? We're ignoring the s^2 term in comparison to the s term in the denominator, or:

$$|\tau_1 \tau_2 \omega^2| \ll |(\tau_1 + \tau_2) \omega| \quad [6-6]$$

This approximation is justified when:

$$\omega \ll \frac{(\tau_1 + \tau_2)}{\tau_1 \tau_2} \ll \frac{1}{\tau_1} + \frac{1}{\tau_2} \quad [6-7]$$

Assume that there is a high-frequency (ω_h) and a low-frequency (ω_l) pole, given by:

$$\begin{aligned} \omega_l &= \frac{1}{\tau_1} \\ \omega_h &= \frac{1}{\tau_2} \end{aligned} \quad [6-8]$$

This approximation says that our estimate is valid if we are interested in frequencies well below the high-frequency pole. This is surely valid if the amplifier response is dominated by the low-frequency pole ω_l .

How accurate is the estimate? Let's take a look at the second-order system transfer function:

$$H(j\omega) = \frac{1}{-\tau_1\tau_2\omega^2 + \cdots + (\tau_1 + \tau_2)(j\omega) + 1} \quad [6-9]$$

Let's assume that the sum of the pole time constants $\tau_1 + \tau_2$ is one second. Therefore, our estimate for -3dB bandwidth of this amplifier is 1 radian per second. We'll rewrite the transfer function using $\omega_{h,est}$ as our estimate for -3dB bandwidth:

$$H(j\omega) = \frac{1}{-\tau_1\tau_2\omega_{h,est}^2 + \cdots + (\tau_1 + \tau_2)(j\omega_{h,est}) + 1} \quad [6-10]$$

Are we justified in neglecting the ω^2 term in the denominator of the polynomial? We note that the magnitude of the ω term is one, since $\omega_{h,est} = 1$ is our estimate for bandwidth. The maximum value of the ω^2 term is 0.25. So, for estimating bandwidth this approximation is justified.

Similarly, we can estimate the bandwidth of a higher-order system with n poles as the reciprocal of the sum of the pole time constants τ_{pn} :

$$H(s) \approx \frac{1}{(\tau_1 + \tau_2 + \cdots + \tau_n)s + 1}$$

$$\omega_h \approx \frac{1}{(\tau_1 + \tau_2 + \cdots + \tau_n)} = \frac{1}{\sum_{pn} \tau_{pn}} \quad [6-11]$$

How do we find the sum of pole time constants $(\tau_1 + \tau_2 + \cdots + \tau_n)$? A proof by Adler at MIT shows that the *sum of the pole time constants* is exactly equal to the *sum of open-circuit time constants*, which is² given as:

$$\tau_1 + \tau_2 + \cdots + \tau_n = \sum_{i=1}^n \tau_{oi} \quad [6-12]$$

The approximate bandwidth of the amplifier is:

$$\omega_h \approx \frac{1}{\sum_{oi} \tau_{oi}} \quad [6-13]$$

This approximation is often surprisingly accurate and the open-circuit time constants are often easy to calculate.....while the pole locations are often very *difficult* to calculate. To calculate each open-circuit time constant τ_{oi} , do the following steps:

- Open circuit all capacitances in the system that contribute to high-frequency bandwidth limitations.
- Find the resistance facing each capacitor's terminals, one-by-one.
- Each individual time constant is found by:

$$\tau_{oi} = R_{oi}C_i \quad [6-14]$$

² Each individual pole time constant is not the same as each open-circuit time constant. However, the *sum* of the pole time constants is equal to the *sum* of the open-circuit time constants.

- Sum the individual open-circuit time constants and find estimate of bandwidth.

$$\omega_h \approx \frac{1}{\sum_{i=1}^n \tau_{oi}} \quad [6-15]$$

Example 6.1: Elementary OCTC example

For the circuit of **Figure 6-2** we'll estimate the -3dB bandwidth using the method of open-circuit time constants, then find the actual -3dB bandwidth by calculation and by using PSPICE.

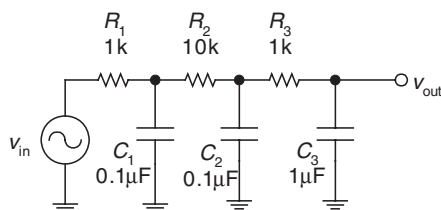


Figure 6-2: Elementary open-circuit time constants example.

There are three capacitors, each of which contributes to bandwidth limitation in this circuit. We'll perform the method of open-circuit time constants using these three bandwidth-limiting capacitances.

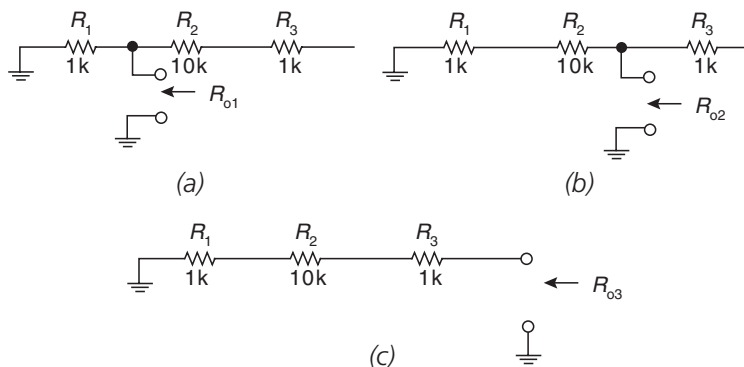


Figure 6-3: Individual circuits for finding open-circuit resistances. (a) For C_1 . (b) For C_2 . (c) For C_3 .

For capacitor C_1 , we open-circuit C_2 and C_3 and find the resistance R_{o1} across the C_1 terminals (**Figure 6-3a**). Note that we also short the input generator, as an input voltage generator provides an incremental short. The open-circuit resistance for C_1 is $1\text{ k}\Omega$ and hence the open-circuit time constant for C_1 , which we'll call τ_{o1} , is 0.1 milliseconds .

For C_2 and C_3 we follow a similar procedure, with the result:

- For C_2 : Open-circuit resistance $R_{o2} = 11 \text{ k}\Omega$, $\tau_{o2} = 1.1 \text{ milliseconds}$
- For C_3 : Open-circuit resistance $R_{o3} = 12 \text{ k}\Omega$, $\tau_{o3} = 12 \text{ milliseconds}$

Our estimate for bandwidth of this circuit is as follows:

$$\begin{aligned}\sum \tau_{oc} &= \tau_{o1} + \tau_{o2} + \tau_{o3} = 1.32 \times 10^{-2} \text{ sec.} \\ \omega_h &\approx \frac{1}{\sum \tau_{oc}} \approx 75.75 \text{ rad/sec} \\ f_h &= \frac{\omega_h}{2\pi} \approx 12 \text{ Hz}\end{aligned} \quad [6-16]$$

PSPICE shows that the -3dB bandwidth $f_h \approx 12 \text{ Hz}$ (**Figure 6-4**). We also see that the dominant bandwidth-limiting capacitor is C_3 , which makes sense since C_3 is larger than the other capacitors.

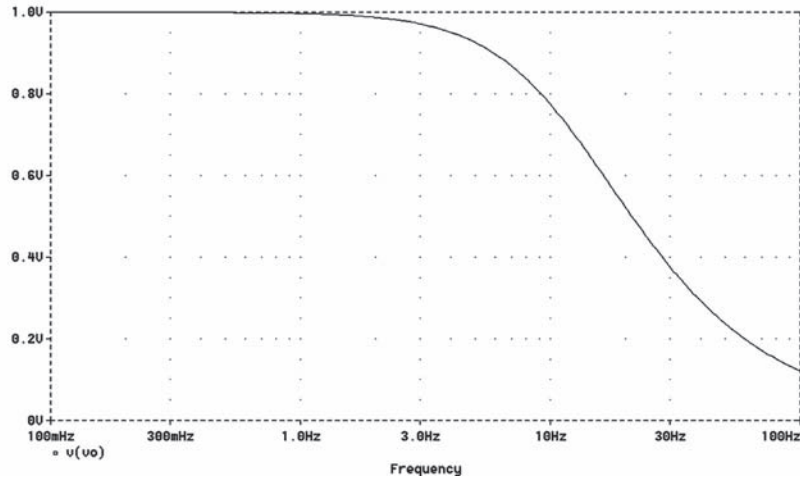


Figure 6-4: SPICE result of AC analysis for simple RC circuit example.

If we want to find the poles and zeros numerically, first we write the node equations. The node equations for this system, in matrix form, are:

$$\begin{bmatrix} -(G_1 + G_2 + C_1 s) & G_2 & 0 \\ G_2 & -(G_2 + G_3 + C_2 s) & G_3 \\ 0 & G_3 & -(G_3 + C_3 s) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_{out} \end{bmatrix} = \begin{bmatrix} -v_{in} G_1 \\ 0 \\ 0 \end{bmatrix} \quad [6-17]$$

A MATLAB script for finding the poles is as follows:

```
function rcrc
% rcrc network for AC&I notes example
R1=1000; G1=1/R1;
R2=10000; G2=1/R2;
R3=1000; G3=1/R3;
C1=10^-7;
C2=10^-7;
C3=10^-6;

G=[-(G1+G2) G2 0;
    G2 -(G2+G3) G3;
    0 G3 -G3];
C=[-C1 0 0;
    0 -C2 0;
    0 0 -C3];

poles=-eig(G/C)
poles_in_Hz=poles/(2*pi)
```

The MATLAB result shows a dominant pole at 12.2 Hz and two other higher frequency poles at 1.65 kHz and 1.99 kHz.

Transistor Amplifier Examples

Example 6.2: Common-emitter amplifier (revisited)

Let's revisit the common-emitter amplifier from the previous chapter (**Figure 6-5**) and estimate its -3dB bandwidth using open-circuit time constants.

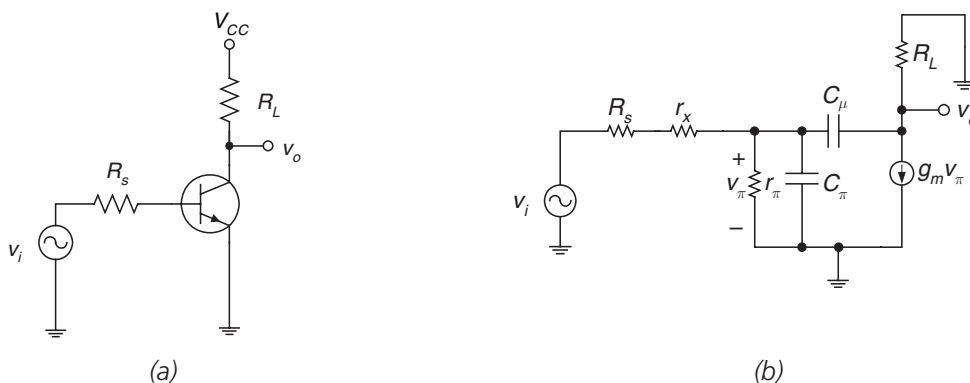


Figure 6-5: Common-emitter amplifier. (a) Circuit, omitting bias details.
(b) Small-signal high-frequency model.

Let's assume that the transistor is biased with a collector current of 2 milliamps and has a small-signal current gain $h_{fe} = 150$ and a collector-base junction capacitance $C_\mu = 2$ picofarads, resulting in the following small-signal parameters:

$$\begin{aligned} g_m &= \frac{|I_C|}{V_{TH}} = \frac{2 \text{ mA}}{26 \text{ mV}} = 0.077 \Omega^{-1} \\ r_\pi &= \frac{h_{fe}}{g_m} = \frac{150}{0.077 \Omega^{-1}} = 1950 \Omega \\ C_\pi &= \frac{g_m}{\omega_T} - C_\mu = \frac{0.077}{2\pi(300 \times 10^6)} - 2 \text{ pF} = 38.8 \text{ pF} \end{aligned} \quad [6-18]$$

Now, we'll estimate the bandwidth using open-circuit time constants and the small high-frequency model of **Figure 6-5**. For C_π , we add a test voltage source v_t and calculate the test current i_t as follows (**Figure 6-6a**). The open-circuit resistance seen across the C_π terminals with C_μ an open circuit is:

$$R_{o\pi} = \frac{v_t}{i_t} \quad [6-19]$$

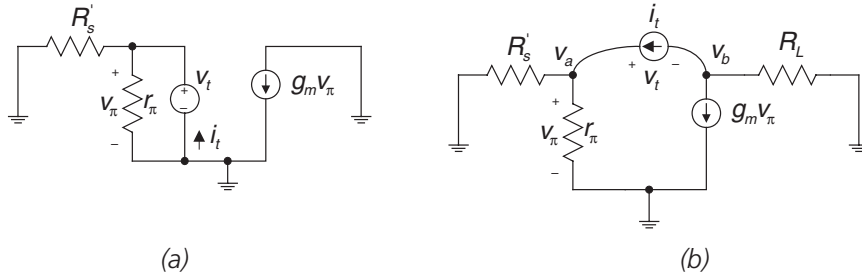


Figure 6-6: Circuits for finding open-circuit resistances.

(a) Circuit for finding C_π time constant of common-emitter amplifier.

(b) Circuit for finding C_μ time constant of common-emitter amplifier.

By inspection, the open-circuit resistance facing C_π is the parallel combination of r_π and the source resistance:

$$R_{o\pi} = r_\pi \parallel R_s' = r_\pi \parallel (R_s + r_x) = 1950 \parallel 1100 = 703 \Omega \quad [6-20]$$

The resultant open-circuit time constant is:

$$\tau_{o\pi} = R_{o\pi} C_\pi = (703 \Omega)(38.8 \text{ pF}) = 27.3 \text{ ns} \quad [6-21]$$

For C_μ , we use the circuit of **Figure 6-6b**. In order to find the test voltage v_t , we find the two node voltages to ground, v_a and v_b . Voltage v_a is easy to find, since the resistance on the lefthand side of the current source is just the open-circuit resistance we found earlier for C_π , or $R_{o\pi}$. The voltage v_b at the right of the current source is:

$$v_b = -(i_t + g_m v_\pi) R_L = -(i_t + g_m i_t R_{o\pi}) R_L \quad [6-22]$$

Hence, the test voltage is:

$$v_t = v_a - v_b = i_t R_{o\pi} - (i_t + g_m i_t R_{o\pi}) R_L \quad [6-23]$$

Solving for v_t/i_t results in:

$$R_{o\mu} = \frac{v_t}{i_t} = R_{o\pi} + (1 + g_m R_{o\pi}) R_L = 5.58 \times 10^4 \Omega \quad [6-24]$$

The resultant open-circuit time constant for C_μ is:

$$\tau_{o\mu} = R_{o\mu} C_\mu = (5.58 \times 10^4 \Omega)(2 \text{ pF}) = 111.7 \text{ ns} \quad [6-25]$$

The sum of the open-circuit time constants is:

$$\sum \tau_{oc} = \tau_{o\pi} + \tau_{o\mu} = 139 \text{ ns} \quad [6-26]$$

Our estimation for –3dB bandwidth is:

$$\omega_h \approx \frac{1}{\sum \tau_{oc}} = 7.19 \times 10^6 \text{ rad/s} \quad [6-27]$$

$$f_h = \frac{\omega_h}{2\pi} = 1.14 \text{ MHz}$$

A comparison of the four solution methods (**Table 6-1**) shows good agreement between the closed-form solution, the one-pole (Miller) approximation, the method of open-circuit time constants and the SPICE simulation (**Figure 6-7**).

Table 6-1: Comparison of results for common-emitter amplifier.

Method	Bandwidth calculation
Closed form (from Chapter 5)	7.2 Mrad/sec (1.15 MHz)
One-pole (Miller) approximation (from Chapter 5)	7.29 Mrad/sec (1.16 MHz)
Open-circuit time constants	7.19 Mrad/sec (1.14 MHz)
SPICE result	7.2 Mrad/sec (~1.15 MHz)

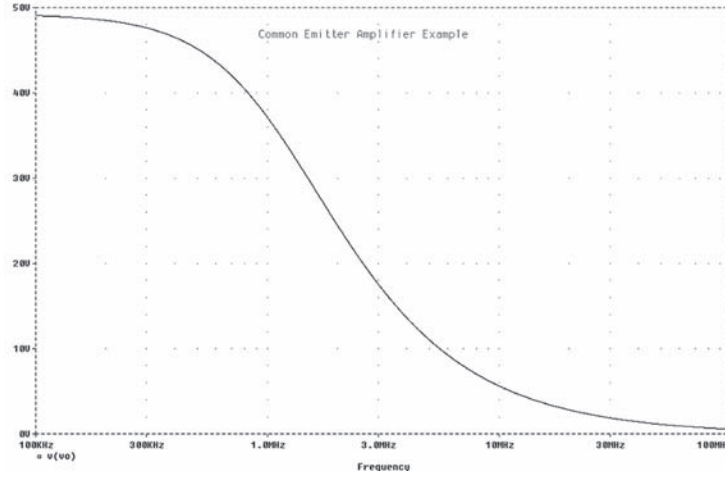


Figure 6-7: SPICE result for common-emitter amplifier.

Using the common-emitter amplifier result as an OCTC sanity check

We're now in a position to use the closed-form result of the previous chapter as a sanity check for the open-circuit time constants method. The closed-form transfer function for the common-emitter amplifier, repeated here for the reader's convenience, is:

$$\frac{v_o(s)}{v_i(s)} = -g_m R_L \frac{G'_s}{G'_s + g_\pi} \left[\frac{1 - \frac{C_\mu}{g_m} s}{\frac{R_L C_\pi C_\mu}{G'_s + g_\pi} s^2 + \frac{1}{G'_s + g_\pi} [R_L (g_m + g_\pi + G'_s) C_\mu + C_\pi + C_\mu] s + 1} \right] \quad [6-28]$$

Note that coefficient of the s term is:

$$\frac{1}{G'_s + g_\pi} [R_L (g_m + g_\pi + G'_s) C_\mu + C_\pi + C_\mu] \quad [6-29]$$

We'll regroup the terms of the " s " coefficient so we can see the terms associated with C_π and C_μ individually.

$$\begin{aligned} & \frac{C_\pi}{G'_s + g_\pi} + \frac{C_\mu}{G'_s + g_\pi} + \frac{1}{G'_s + g_\pi} [R_L (g_m + g_\pi + G'_s) C_\mu] \\ & \approx \frac{C_\pi}{G'_s + g_\pi} + \left(\frac{1 + g_m R_L}{G'_s + g_\pi} \right) C_\mu + R_L C_\mu \end{aligned} \quad [6-30]$$

The OCTC method allows us to calculate the coefficient of the s term, so the OCTC result should give us the same result as the closed-form equation above. The individual open-circuit time constants for C_π and C_μ are:

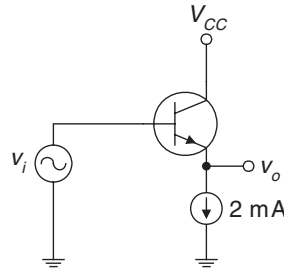
$$\begin{aligned}\tau_{o\pi} &= (r_\pi \parallel R'_s) C_\pi = \frac{C_\pi}{g_\pi + G'_s} \\ \tau_{o\mu} &= (r_\pi \parallel R'_s + R_L + g_m (r_\pi \parallel R'_s) R_L) C_\mu = \left(\frac{1 + g_m R_L}{G'_s + g_\pi} + R_L \right) C_\mu = \left(\frac{1 + g_m R_L}{G'_s + g_\pi} \right) C_\mu + R_L C_\mu\end{aligned}\quad [6-31]$$

Note that the two methods give the same answer for the coefficient of the s term, as expected.

Example 6.3: Emitter-follower bandwidth estimate

Using open-circuit time constants, let's estimate the bandwidth of the emitter-follower (**Figure 6-8**) from the previous chapter.

Figure 6-8: Emitter-follower.

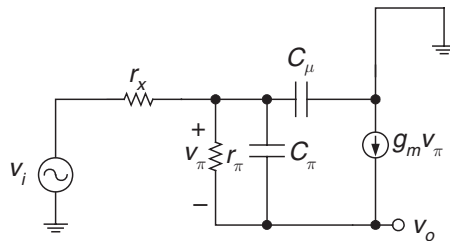


The emitter-follower is biased with a 2-milliamp current source, resulting in the following small-signal parameters:

$$\begin{aligned}g_m &= \frac{|I_C|}{V_{TH}} = \frac{2 \text{ mA}}{26 \text{ mV}} = 0.077 \Omega^{-1} \\ r_\pi &= \frac{h_{fe}}{g_m} = \frac{150}{0.077 \Omega^{-1}} = 1950 \Omega \\ C_\pi &= \frac{g_m}{\omega_T} - C_\mu = \frac{0.077}{2\pi(300 \times 10^6)} - 2 \text{ pF} = 38.8 \text{ pF}\end{aligned}\quad [6-32]$$

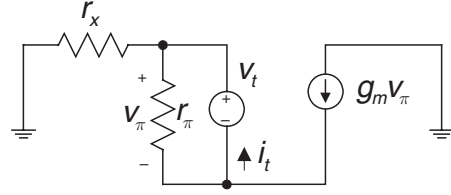
The small-signal model of the emitter-follower is shown in **Figure 6-9**.

Figure 6-9: Emitter-follower small-signal model.



Next, we'll estimate the bandwidth using open-circuit time constants. For C_π , add a test voltage source v_t and measure the test current i_t as shown in **Figure 6-10**.

Figure 6-10: Circuit for finding C_π time constant for emitter-follower.



The test current is:

$$i_t = \frac{v_t}{r_\pi} + g_m v_t \quad [6-33]$$

Hence, the open-circuit resistance facing C_π is (noting that $g_m \gg 1/r_\pi$):

$$R_{o\pi} = \frac{v_t}{i_t} = \frac{1}{g_m + g_\pi} \approx \frac{1}{g_m} \approx 13.07\Omega \quad [6-34]$$

and the resultant time constant is:

$$\tau_{o\pi} = R_{o\pi} C_\pi = (13.07\Omega)(38.8 \text{ pF}) = 0.51 \text{ ns} \quad [6-35]$$

For C_μ , we use the circuit of **Figure 6-11a**.

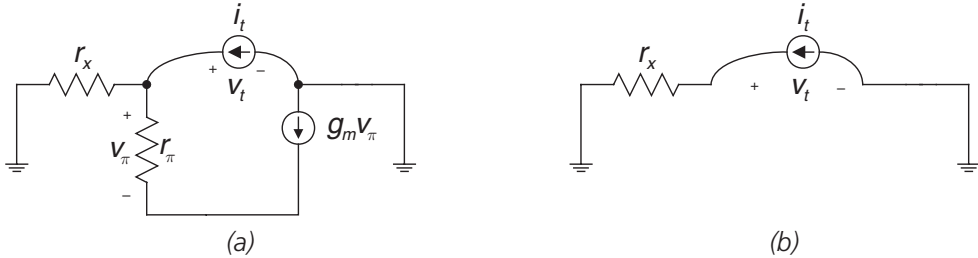


Figure 6-11: Circuit for finding C_μ time constant for emitter-follower.
(a) Original circuit. (b) Simplified circuit acknowledging the fact that $g_m v_\pi = 0$.

Now, we can make an important simplification that greatly eases the analysis for this case. Note that there is a current loop containing r_π and the $g_m v_\pi$ dependent generator. If we sum currents at the emitter node, the following holds:

$$g_m v_\pi = -\frac{v_\pi}{r_\pi} \quad [6-36]$$

Recognizing that this can only be true when $v_\pi = 0$, we note that the current generator must have zero current and hence we can eliminate it, resulting in the simplified circuit of **Figure 6-11b**. Therefore, the open-circuit resistance for C_μ is:

$$R_{o\mu} = r_x = 100\Omega \quad [6-37]$$

and the resultant time constant is:

$$\tau_{o\mu} = R_{o\mu} C_{\mu} = (100\Omega)(2 \text{ pF}) = 0.2 \text{ ns} \quad [6-38]$$

The sum of the open-circuit time constants is:

$$\sum \tau_{oc} = \tau_{o\pi} + \tau_{o\mu} = 0.71 \text{ ns} \quad [6-39]$$

and our estimation for -3dB bandwidth is:

$$\omega_h \approx \frac{1}{\sum \tau_{oc}} = 1.41 \times 10^9 \text{ rad/s} \quad [6-40]$$

$$f_h = \frac{\omega_h}{2\pi} = 224 \text{ MHz}$$

A SPICE result showing the gain and bandwidth is shown in **Figure 6-12**. This result shows that the gain is approximately unity (as expected) but the -3dB bandwidth is approximately 800 MHz. What is going on here?

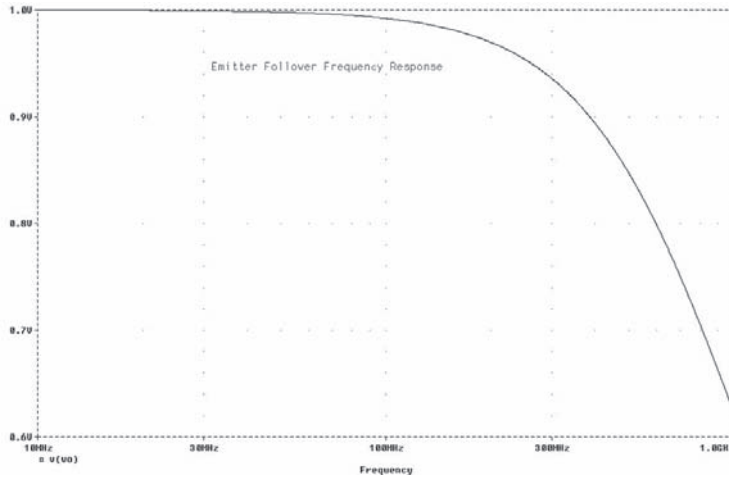


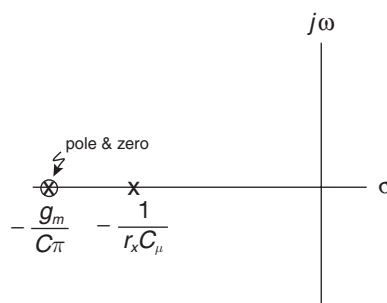
Figure 6-12: SPICE result for emitter-follower.

It's time to look at the closed-form solution from the previous chapter in more detail. Previously, we found the transfer function of the emitter-follower to be:³

$$\frac{v_o(s)}{v_i(s)} = \frac{\frac{C_{\pi}}{g_m} s + 1}{\frac{r_x C_{\pi} C_{\mu}}{g_m} s^2 + \left[r_x C_{\mu} + \frac{C_{\pi}}{g_m} \right] s + 1} \quad [6-41]$$

³ Note that the sum of open-circuit time constants for the emitter-follower equals the coefficient of the s term in the closed-form solution.

Figure 6-13: Emitter-follower pole-zero plot.



There are two poles and a zero, and the zero nearly⁴ cancels the high-frequency pole. The method of open-circuit time constants does not account for this zero, and hence greatly underestimates the bandwidth in this case.

Furthermore, the preceding result must be taken with a grain of salt, as the model predicts a bandwidth in excess of the unity current-gain bandwidth product f_T of the transistor. These are caveats to keep in mind when using the open-circuit time constants method.

Example 6.4: Differential amplifier

A differential amplifier is shown in **Figure 6-14a**. For differential-mode excitation, the small-signal model of **Figure 6-14b** is valid. We can use this model and the method of open-circuit time constants to estimate the bandwidth of this amplifier. Note that the differential half-circuit is identical to the common-emitter amplifier we worked with previously. Hence, the circuits for finding the open-circuit time constants (**Figure 6-15**) are the same as the common-emitter amplifier.

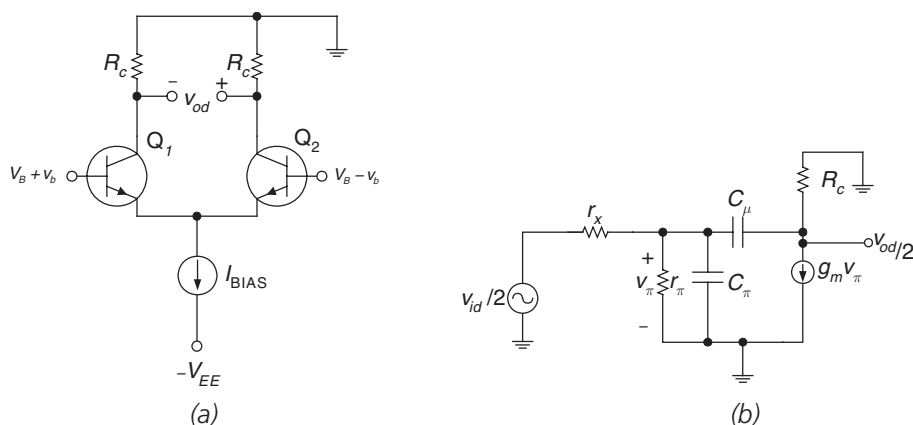


Figure 6-14: Differential amplifier. (a) Circuit. (b) Differential-mode half circuit.

⁴ The pole and zero nearly cancel out; in the closed-form solution we made an approximation or two, and the pole and zero don't exactly cancel.

Chapter 6

Let's find the differential gain and bandwidth of a differential amplifier with the following parameters:

- $R_C = 3.3\text{k}$
- $I_{\text{BIAS}} = 4$ milliamps
- $h_{fe} = 100$
- $f_T = 300$ MHz
- $r_x = 100\Omega$
- $C_\mu = 1$ pF.

The small-signal parameters for this amplifier are:

$$\begin{aligned}
 g_m &= \frac{|I_C|}{V_{TH}} = \frac{2 \text{ mA}}{26 \text{ mV}} = 0.077 \Omega^{-1} \\
 r_\pi &= \frac{h_{fe}}{g_m} = \frac{100}{0.077 \Omega^{-1}} = 1299 \Omega \\
 C_\pi &= \frac{g_m}{\omega_T} - C_\mu = \frac{0.077}{2\pi(300 \times 10^6)} - 2 \text{ pF} = 39 \text{ pF}
 \end{aligned} \tag{6-42}$$

The small-signal differential mode gain is:

$$\frac{v_{od}}{v_{id}} = -g_m R_L \left(\frac{r_\pi}{r_x + r_\pi} \right) \approx -236 \tag{6-43}$$

Our open-circuit time constant calculations are as follows. For C_π , we use the circuit of **Figure 6-15a**:

$$\begin{aligned}
 R_{o\pi} &= r_x \parallel r_\pi \approx r_x \approx 100 \Omega \\
 \tau_{o\pi} &= R_{o\pi} C_\pi = (100)(39 \text{ pF}) = 3.9 \text{ ns}
 \end{aligned} \tag{6-44}$$

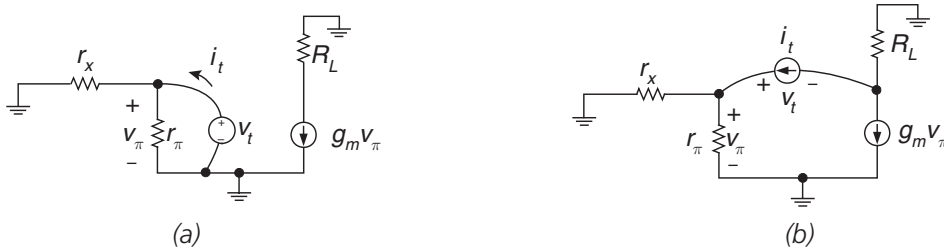


Figure 6-15: Circuits for finding OCTCs for differential amplifier.

(a) Circuit for C_π . (b) Circuit for C_μ .

For C_μ , we use the circuit of **Figure 6-15b** and we find:

$$R_{o\mu} = r_x + R_L + g_m r_x R_L = 100 + 3300 + (0.077)(100)(3300) = 28810\Omega \quad [6-45]$$

$$\tau_{o\mu} = R_{o\mu} C_\mu = (28810)(2 \text{ pF}) = 57.6 \text{ ns}$$

The sum of open-circuit time constants for this amplifier is 61.5 nanoseconds, and our estimate for bandwidth is:

$$\omega_h \approx \frac{1}{61.5 \text{ ns}} \approx 16.25 \text{ Mrad/sec} \quad [6-46]$$

$$f_h = \frac{\omega_h}{2\pi} \approx 2.58 \text{ MHz}$$

PSPICE (**Figure 6-16**) shows that the gain is -236 as expected and the -3dB bandwidth is $\sim 2.8 \text{ MHz}$, which is a little higher than the bandwidth predicted by open-circuit time constants.

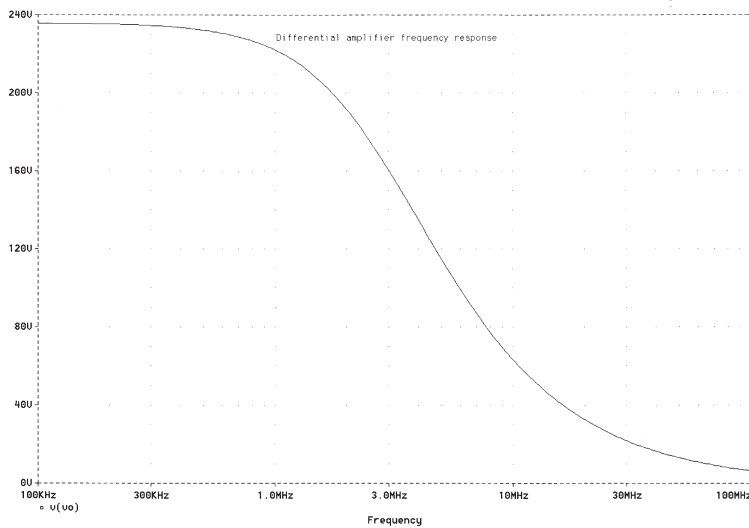
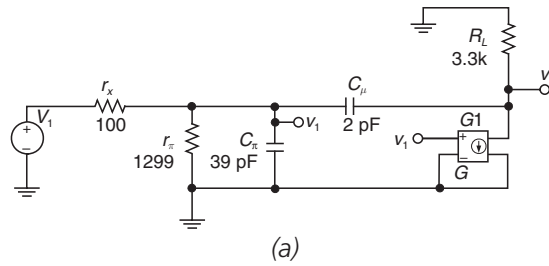


Figure 6-16: PSPICE result for common-base amplifier.
a) PSPICE circuit. (b) Frequency response of common-base amplifier showing low frequency gain of -236 and -3dB bandwidth of 2.8 MHz .

Example 6.5: Design case study using open-circuit time constants

In the previous examples, we used open-circuit time constants to analyze the bandwidth of some circuits. In the following example, we'll use open-circuit time constants as a tool in an iterative design exercise. Let's design a transistor amplifier to meet the following gain and bandwidth specifications:

- Magnitude of midband gain: $|A_v| : > 100$
- Input source resistance: $R_s = 2 \text{ k}\Omega$ (high impedance source, such as a microphone)
- Load capacitance: $C_L = 10 \text{ pF}$ (scope probe, stray capacitance, etc.)
- -3dB bandwidth: $f_h > 10 \text{ MHz}$
- Transistor parameters: $h_{fe} = 150$; $f_T = 300 \text{ MHz}$; $C_\mu \approx 2 \text{ pF}$

As an initial guess at a circuit topology, let's consider the AC-coupled common-emitter amplifier of **Figure 6-17a**. The Bode plot of this amplifier looks qualitatively like that in **Figure 6-17b**. The high-frequency rolloff is due to the bandwidth-limiting effects of transistor internal capacitances C_π and C_μ . The low-frequency rolloff is due to the coupling capacitor C_C and the emitter bypass capacitor C_E .⁵ In the next chapter, we'll estimate the low-frequency breakpoint f_L using the method of *short-circuit time constants*. In this section, we'll estimate the high-frequency rolloff f_h using open-circuit time constants.

TRY #1: Common-emitter amplifier

First, let's arbitrarily assume that the collector current of the transistor is 2 milliamps.⁶ Other small-signal parameters are:

$$\begin{aligned}
 g_{m1} &= \frac{|I_{C1}|}{V_{TH}} = \frac{2 \text{ mA}}{26 \text{ mV}} = 0.077 \Omega^{-1} \\
 r_{\pi1} &= \frac{h_{fe1}}{g_{m1}} = \frac{150}{0.077 \Omega^{-1}} = 1948 \Omega \\
 C_{\pi1} &= \frac{g_{m1}}{\omega_T} - C_{\mu1} = \frac{0.077}{2\pi(300 \times 10^6)} - 2 \text{ pF} = 39 \text{ pF}
 \end{aligned} \tag{6-47}$$

⁵ We'll assume that C_C and C_E are so large that the low-frequency breakpoint f_L is well below the high-frequency breakpoint f_h .

⁶ We can easily bias this transistor at a collector current of 2 milliamps by carefully designing the values of R_{B1} , R_{B2} and R_E as shown in the previous discussions on transistor biasing.

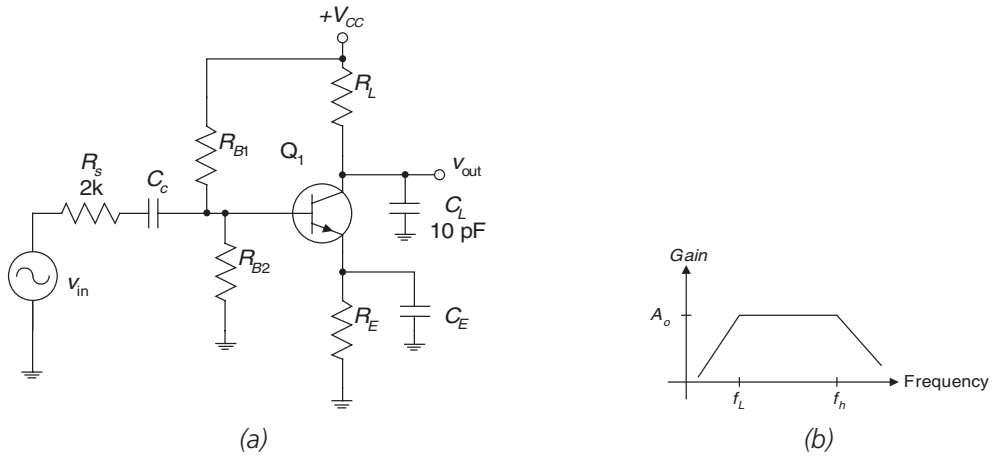


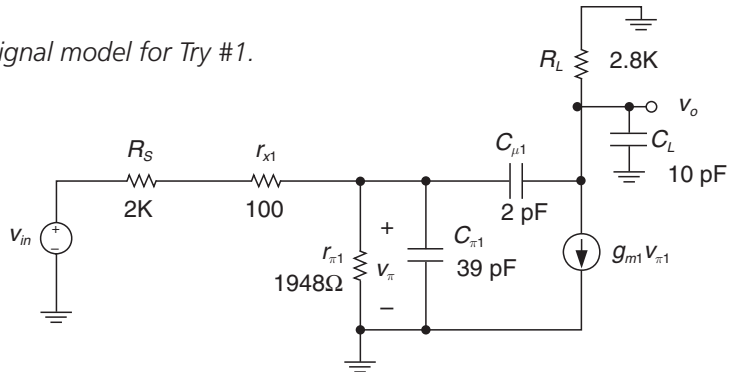
Figure 6-17: Initial guess at topology (Try #1).
(a) Circuit. (b) Qualitative view of gain curve for amplifier.

In order to find the midband gain, we use the small-signal model of **Figure 6-18**.⁷ We find that the small-signal low-frequency gain is:

$$A_v = -g_{m1}R_L \left(\frac{r_{\pi1}}{r_{\pi} + r_{x1} + R_s} \right) \quad [6-48]$$

Note that $-g_m R_L$ is what the gain would be if the source resistor (R_s) and transistor base spreading resistance (r_x) were zero. Using this gain equation, we find that the minimum value of R_L that achieves our necessary gain of 100 is $R_L = 2.7 \text{ k}\Omega$. We'll use $R_L = 2.8 \text{ k}\Omega$ to ensure a little bit of extra gain.

Figure 6-18: Small-signal model for Try #1.



⁷ I have assumed that R_{B1} and R_{B2} are so big that they don't affect the gain and bandwidth calculations significantly. In order for this to be true, R_{B1} and R_{B2} need to be large compared to R_s , r_x and r_{π} . If this is not the case, we need to include R_{B1} and R_{B2} in the gain and bandwidth calculations. We also assume that we're operating at frequencies high enough so that the coupling and emitter bypass capacitors act as short circuits.

Chapter 6

Next, in order to estimate the bandwidth we apply the method of open-circuit time constants. In order to find the open-circuit resistance across $C_{\pi 1}$, we use the circuit of **Figure 6-19a**. The open-circuit resistance and time constant are:

$$R_{\pi 1} = r_{\pi 1} \parallel (R_s + r_{x1}) = 1948 \parallel 2100 = 1010 \Omega$$

$$\tau_{\pi 1} = R_{\pi 1} C_{\pi 1} = (1010)(39 \text{ pF}) = 39 \text{ ns}$$
[6-49]

To find the open-circuit resistance for $C_{\mu 1}$, we use the circuit of **Figure 6-19b**. We find that the open-circuit resistance across the $C_{\mu 1}$ terminals is the impedance to the left ($R_{\pi 1}$), the impedance to the right (R_L) and a third term due to the $g_m v_{\pi}$ dependent generator. The result is:

$$R_{\mu 1} = R_{\pi 1} + R_L + g_{m1} R_{\pi 1} R_L = 1010 + 2700 + (0.077)(1010)(2800) = 221.5 \text{ k}\Omega$$

$$\tau_{\mu 1} = R_{\mu 1} C_{\mu 1} = (221.5 \text{ k}\Omega)(2 \text{ pF}) = 443 \text{ ns}$$
[6-50]

The open-circuit resistance for the output load capacitor (**Figure 6-19c**) is easy to find by inspection, because the dependent current source is disabled. The result is:

$$R_{o,L} = R_L = 2.8 \text{ k}\Omega$$

$$\tau_{LOAD} = R_{o,L} C_L = (2800)(10 \text{ pF}) = 28 \text{ ns}$$
[6-51]

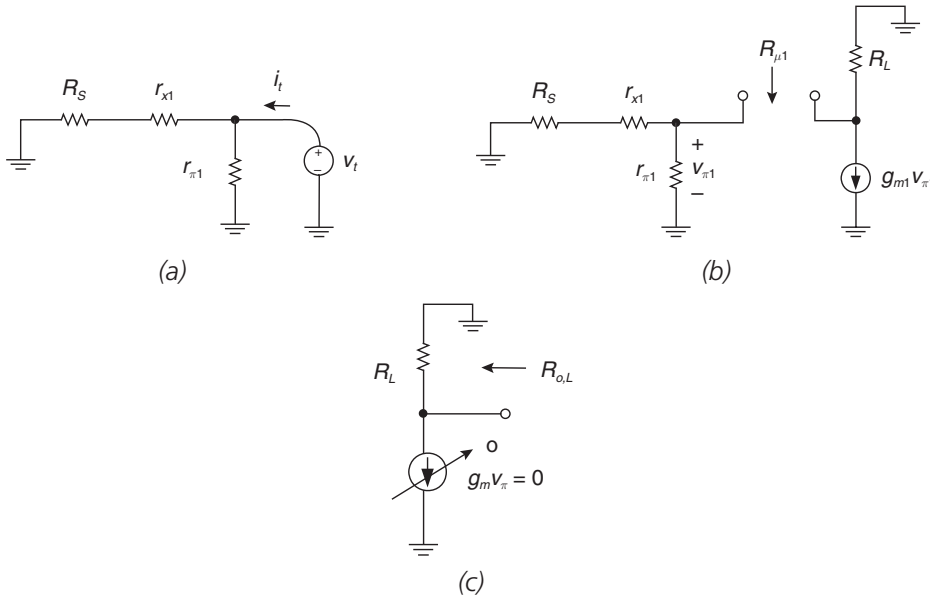


Figure 6-19: Circuits for determining open-circuit resistances. (a) OCTC circuit for $C_{\pi 1}$. (b) OCTC circuit for $C_{\mu 1}$. (c) OCTC circuit for load capacitor C_{LOAD} .

The sum of open-circuit time constants for this amplifier is 510 nanoseconds, and our estimate of bandwidth is:

$$\omega_h \approx \frac{1}{510 \text{ ns}} \approx 1.96 \text{ Mrad/sec}$$

$$f_h = \frac{\omega_h}{2\pi} \approx 312 \text{ kHz}$$
[6-52]

PSPICE (**Figure 6-20**) shows that the gain is ~ -103.8 as calculated and the bandwidth is approximately 314 kHz. A summary of the open-circuit time constant results for Try #1 is shown in **Table 6-2**. There is one dominant pole in the amplifier (due to the Miller effect), so in this case the open-circuit time constants estimate is pretty good. We haven't met the bandwidth specification, so we need to iterate the design.

Table 6-2: OCTC summary for Try #1.

	TRY #1
$\tau_{\pi 1}$	39 ns
$\tau_{\mu 1}$	443 ns
τ_{LOAD}	28 ns
$\sum \tau_{\text{oc}}$	510 ns
$\omega_{h, \text{est}}$	1.96 MR/s
$f_{h, \text{est}}$	312 kHz
f_h from PSPICE	314 kHz

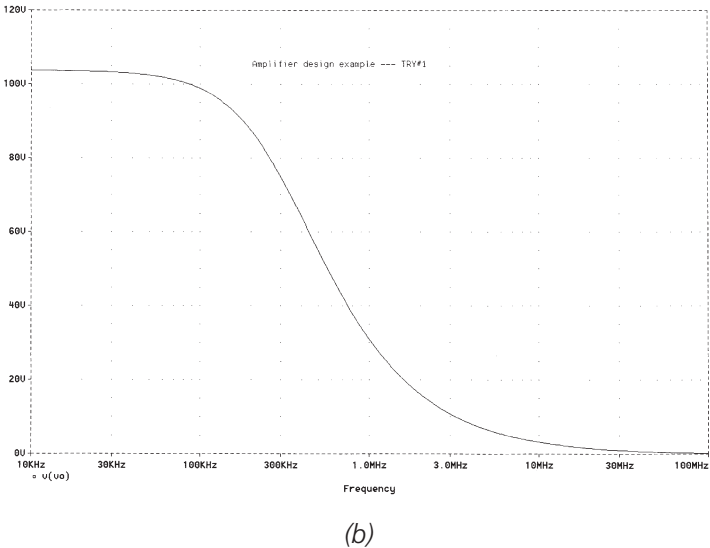
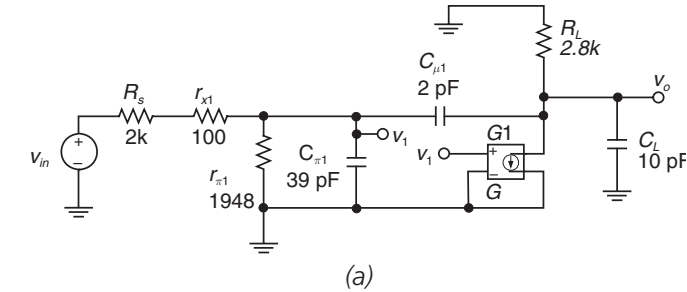


Figure 6-20: Frequency response for Try #1. (a) PSPICE circuit, omitting biasing details. (b) PSPICE frequency-response result.

TRY #2—Common-emitter + cascode

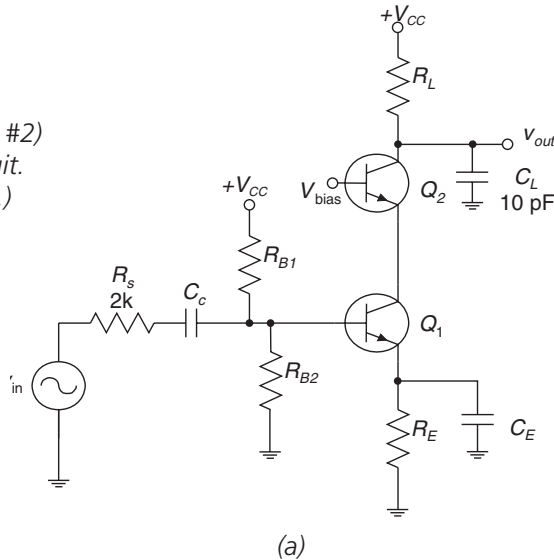
In Try #1, the open-circuit time constants method identifies the time constant of $C_{\mu 1}$ as the dominant bandwidth bottleneck for this amplifier. This result is not surprising considering that we are asking a single common-emitter gain stage to provide all of the gain, and that the Miller effect has a major effect in high-gain common-emitter amplifiers. In fact, a *very* rough estimate of bandwidth due to Miller effect (without resorting to the full Miller approximation given before) is:

$$\omega_h \approx \frac{1}{R_s (AC_{\mu 1})} \approx \frac{1}{(2000)(100)(2 \times 10^{-12})} \approx 2.5 \text{ MR/s}$$

$$f_h \approx \frac{\omega_h}{2\pi} \approx 398 \text{ kHz}$$
[6-53]

In order to reduce the effects of the Miller effect, we can reduce the voltage gain of transistor Q_1 by using to a *cascode*⁸ transistor, as in **Figure 6-21a**. Transistor Q_1 generates a controlled output current that is buffered by common-base transistor Q_2 . The resistance looking into the emitter of Q_2 is very low, so the voltage gain at the collector of Q_1 is low, hence reducing the Miller effect. Transistor Q_2 provides approximately unity current gain; the load resistor at the collector of Q_2 converts the current to an output voltage. We'll assume that the small-signal parameters for Q_1 and Q_2 are identical,⁹ resulting in the small-signal model of **Figure 6-21b**.

Figure 6-21: Second attempt (Try #2)
with cascode amplifier. (a) Circuit.
(Continued on following page.)



⁸ The term *cascode* dates back to vacuum tube circuits. See e.g., F.V. Hunt and R.W. Hickman, “On Electronic Voltage Stabilizers,” *Review of Scientific Instruments*, vol. 10, Jan. 1939, pp. 6–21.

⁹ The two transistors operate at the same collector currents so g_m , r_π and r_x are the same for both. We should actually double-check the values of $C_{\mu 1}$ and $C_{\mu 2}$. With the cascode transistor, Q_1 now operates at a lower V_{CB} than in Try #1, so $C_{\mu 1}$ may go up a little bit. In order to be more rigorous, we should find the V_{CB} of each transistor, then find the C_μ for each transistor individually. However, for this design example let's assume that $C_{\mu 1} = C_{\mu 2} = 2 \text{ pF}$.

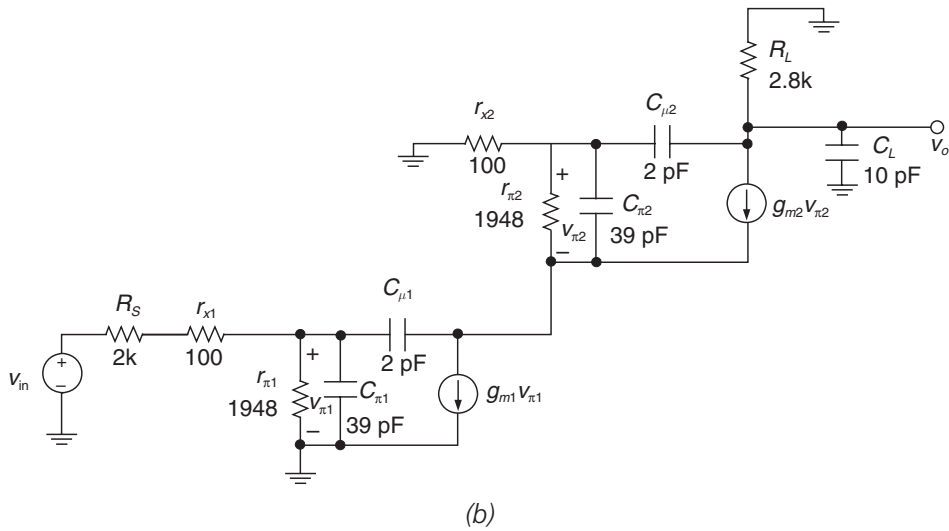


Figure 6-21 (continued): (b) Small-signal model.

We'll now find the five open-circuit time constants for this cascode amplifier. Note that the OCTC for $C_{\pi 1}$ hasn't changed since Try #1; it is still 39 nanoseconds. The open-circuit time constant for $C_{\mu 1}$ requires a little bit more work. In **Figure 6-22a** we see the OCTC circuit for $C_{\mu 1}$. We have replaced Q_2 in this circuit with the output resistance r_{out2} seen at the emitter of Q_2 . This output resistance at the emitter of Q_2 is:¹⁰

$$r_{out2} = \frac{r_{x2} + r_{\pi 2}}{1 + h_{fe2}} = \frac{100 + 1948}{151} = 13.6\Omega \quad [6-54]$$

We now find the OCTC as follows:

$$\begin{aligned} R_{\mu 1} &= R_{\pi 1} + r_{out2} + g_m R_{\pi 1} r_{out2} = 1010 + 13.6 + (0.077)(1010)(13.6) = 2081\Omega \\ \tau_{\mu 1} &= R_{\mu 1} C_{\mu 1} = (2081)(2 \text{ pF}) = 4.2 \text{ ns} \end{aligned} \quad [6-55]$$

Note that we have drastically reduced this open-circuit time constant from Try #1, since we have killed the Miller effect. In fact, we have reduced this OCTC by about a factor of 100, consistent with our understanding of the Miller effect.

For $C_{\pi 2}$, we use the circuit of **Figure 6-22b**. Note that the impedance looking into the collector of Q_1 is very high, so the emitter of Q_2 looks into an open circuit. In order to find the OCTC for $C_{\pi 2}$, we add a voltage source and find the test current as follows:

$$i_t = \frac{v_t}{r_{\pi 2}} + g_{m2} v_t \approx g_{m2} v_t \quad [6-56]$$

¹⁰ We can use the results of the previous chapter where we found the output resistance of the emitter-follower.

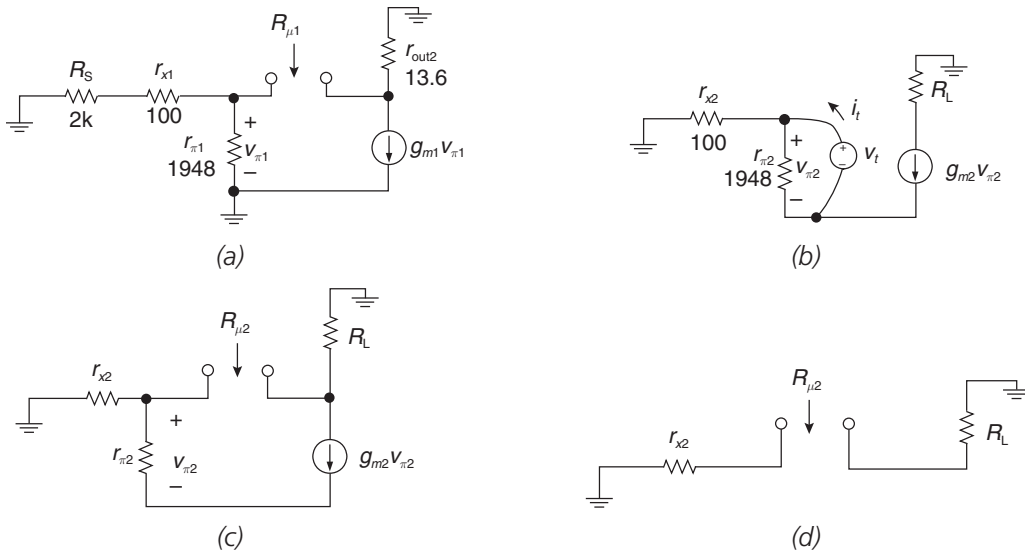


Figure 6-22: Incremental circuits for finding OCTCs for Try #2.
 (a) $C_{\mu 1}$. (b) $C_{\pi 2}$. (c) $C_{\mu 2}$. (d) Simplified circuit for $C_{\mu 2}$.

The open-circuit resistance and open-circuit time constant for $C_{\pi 2}$ are:

$$R_{\pi 2} = \frac{v_t}{i_t} \approx \frac{1}{g_{m2}} \approx 13 \Omega \quad [6-57]$$

$$\tau_{\pi 2} = R_{\pi 2} C_{\pi 2} = (13)(39 \text{ pF}) = 0.5 \text{ ns}$$

For $C_{\mu 2}$, we use the circuit of **Figure 6-22c**. At first glance this circuit looks like a horrible mess, but a little thought is in order before writing node equations. Note that since the output resistance of Q_1 is so high, all the current from the $g_{m2} v_{\pi 2}$ generator flows through $r_{\pi 2}$. This in turn constrains the following to be true:

$$v_{\pi 2} = -g_{m2} r_{\pi 2} v_{\pi 2} \quad [6-58]$$

This can only be true if $v_{\pi 2} = 0$, so we can further simplify by noting that with $v_{\pi 2} = 0$, the $g_{m2} v_{\pi 2}$ current generator is also equal to zero. This results in the simplified circuit of **Figure 6-22d**. We can now find the open-circuit resistance by inspection and time constant for $C_{\mu 2}$.

$$R_{\mu 2} = r_{x2} + R_L = 2900 \Omega \quad [6-59]$$

$$\tau_{\mu 2} = R_{\mu 2} C_{\mu 2} (2900)(2 \text{ pF}) = 5.8 \text{ ns}$$

The open-circuit time constant for the load capacitor is still 28 nanoseconds. Our sum of open-circuit time constants is now 77.5 nanoseconds, and our estimate of bandwidth is:

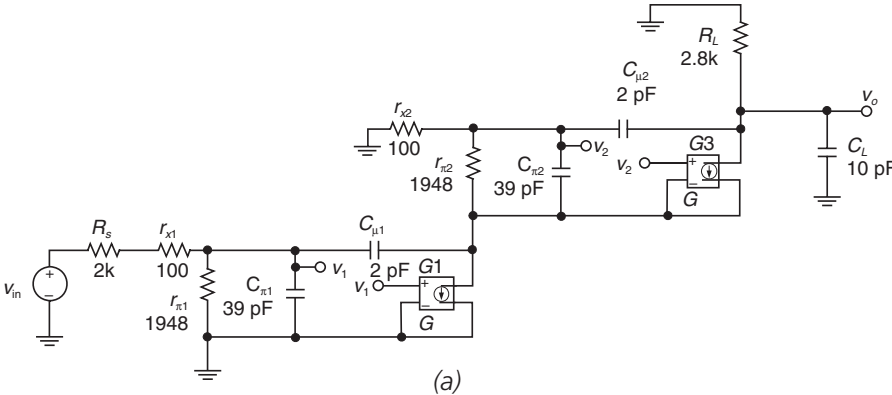
$$\omega_h \approx \frac{1}{77.5 \text{ ns}} \approx 12.9 \text{ Mrad/sec} \quad [6-60]$$

$$f_h \frac{\omega_h}{2\pi} \approx 2.05 \text{ MHz}$$

PSPICE (**Figure 6-23**) shows that the gain is ~ -103.1 as calculated and the bandwidth is approximately 2.68 MHz. A summary of the open-circuit time constant results for Try #2 is shown in **Table 6-3**. We've improved the overall bandwidth significantly by adding the cas-code, but we haven't yet met the bandwidth specification. We'll iterate the design again.

Table 6-3: OCTC summary for Try #1 and #2.

	TRY #1	TRY #2
$\tau_{\pi 1}$	39 ns	39 ns
$\tau_{\mu 1}$	443 ns	4.2 ns
$\tau_{\pi 2}$	---	0.5 ns
$\tau_{\mu 2}$	---	5.8 ns
τ_{LOAD}	28 ns	28 ns
$\sum \tau_{\text{OC}}$	510 ns	77.5 ns
$\omega_{h,\text{est}}$	1.96 MR/s	12.9 MR/s
$f_{h,\text{est}}$	312 kHz	2.05 MHz
f_h from PSPICE	314 kHz	2.68 MHz



(b)

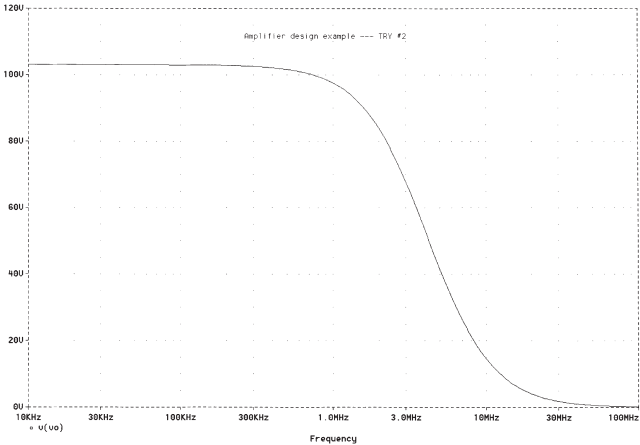


Figure 6-23: Frequency response for Try #2. (a) Circuit. (b) PSPICE frequency-response result.

TRY #3—Emitter-follower + common-emitter + cascode

After Try #2, we see that the dominant time constant is due to $C_{\pi 1}$ of Q_1 . This is due to the large source resistance R_s interacting with $C_{\pi 1}$. We can reduce this effect by buffering the source resistance from the common-emitter transistor with a unity gain emitter-follower buffer, as shown in **Figure 6-24a**. Q_3 is an emitter-follower that has high input impedance and low output impedance. For simplicity, let's assume that Q_3 is biased at the same collector current as the other transistors and has the same small-signal parameters as Q_1 and Q_2 . The small-signal model for Try #3 is shown in **Figure 6-24b**.

The addition of Q_3 and its high input impedance also reduces the gain loading effect due to the high source impedance R_s . Therefore, to achieve our desired gain of 100 we can reduce the load resistor R_L to approximately 1.5 k Ω . As we'll see later on, this also helps with improving the bandwidth.

The output impedance seen at the emitter of Q_3 is:

$$r_{out,Q3} \approx \frac{R_s + r_{x3} + r_{\pi 3}}{1 + h_{fe3}} = \frac{2000 + 100 + 1948}{151} = 27\Omega \quad [6-61]$$

We can now work on the open-circuit time constants for Q_1 . For $C_{\pi 1}$, the calculation is as follows:

$$\begin{aligned} R_{\pi 1} &= r_{\pi 1} \parallel (r_{out,Q3} + r_{x1}) = 1948 \parallel 127 = 119\Omega \\ \tau_{\pi 1} &= R_{\pi 1} C_{\pi 1} = (119)(39 \text{ pF}) = 4.6 \text{ ns} \end{aligned} \quad [6-62]$$

For $C_{\mu 1}$ we find:

$$\begin{aligned} R_{\mu 1} &= R_{\pi 1} + r_{out2} + g_{m2} R_{\pi 1} r_{out2} = 119 + 13.6 + (0.077)(119)(13.6) = 257\Omega \\ \tau_{\mu 1} &= R_{\mu 1} C_{\mu 1} = (257)(2 \text{ pF}) = 0.5 \text{ ns} \end{aligned} \quad [6-63]$$

For transistor Q_2 , the open-circuit time constant for $C_{\pi 2}$ is unchanged at 0.5 nanoseconds. The open-circuit time constant for $C_{\mu 2}$ is reduced from Try #2, since we have lowered the value of the load resistor:

$$\begin{aligned} R_{\mu 2} &= r_{x2} + R_L = 1600\Omega \\ \tau_{\mu 2} &= R_{\mu 2} C_{\mu 2} = (1600)(2 \text{ pF}) = 3.2 \text{ ns} \end{aligned} \quad [6-64]$$

For the emitter-follower transistor Q_3 , we find for $C_{\pi 3}$:

$$\begin{aligned} R_{\pi 3} &= \frac{1}{g_{m3}} \approx 13\Omega \\ \tau_{\pi 3} &= R_{\pi 3} C_{\pi 3} = (13)(39 \text{ pF}) = 0.5 \text{ ns} \end{aligned} \quad [6-65]$$

Next, for $C_{\mu 3}$ we find:

$$\begin{aligned} R_{\mu 3} &= [r_{x3} + R_s] \parallel [r_{\pi 3} + (1 + h_{fe3})(r_{x1} + r_{\pi 1})] \approx r_{x3} + R_s = 2100\Omega \\ \tau_{\mu 3} &= R_{\mu 3} C_{\mu 3} = (2100)(2 \text{ pF}) = 4.2 \text{ ns} \end{aligned} \quad [6-66]$$

The open-circuit time constant due to the load capacitor is 15 nanoseconds. The sum of open-circuit time constants for this entire circuit is now 28.5 nanoseconds, and our estimate of bandwidth is:

$$\omega_h \approx \frac{1}{28.5 \text{ ns}} \approx 35.1 \text{ Mrad/sec}$$

[6-67]

$$f_h = \frac{\omega_h}{2\pi} \approx 5.6 \text{ MHz}$$

PSPICE (**Figure 6-25**) shows that the gain is ~ -107.7 and the bandwidth is approximately 9.4 MHz. A summary of the open-circuit time constant results for Try #3 is shown in **Table 6-4**. We still haven't quite met the 10-MHz bandwidth specification, so we'll iterate again.

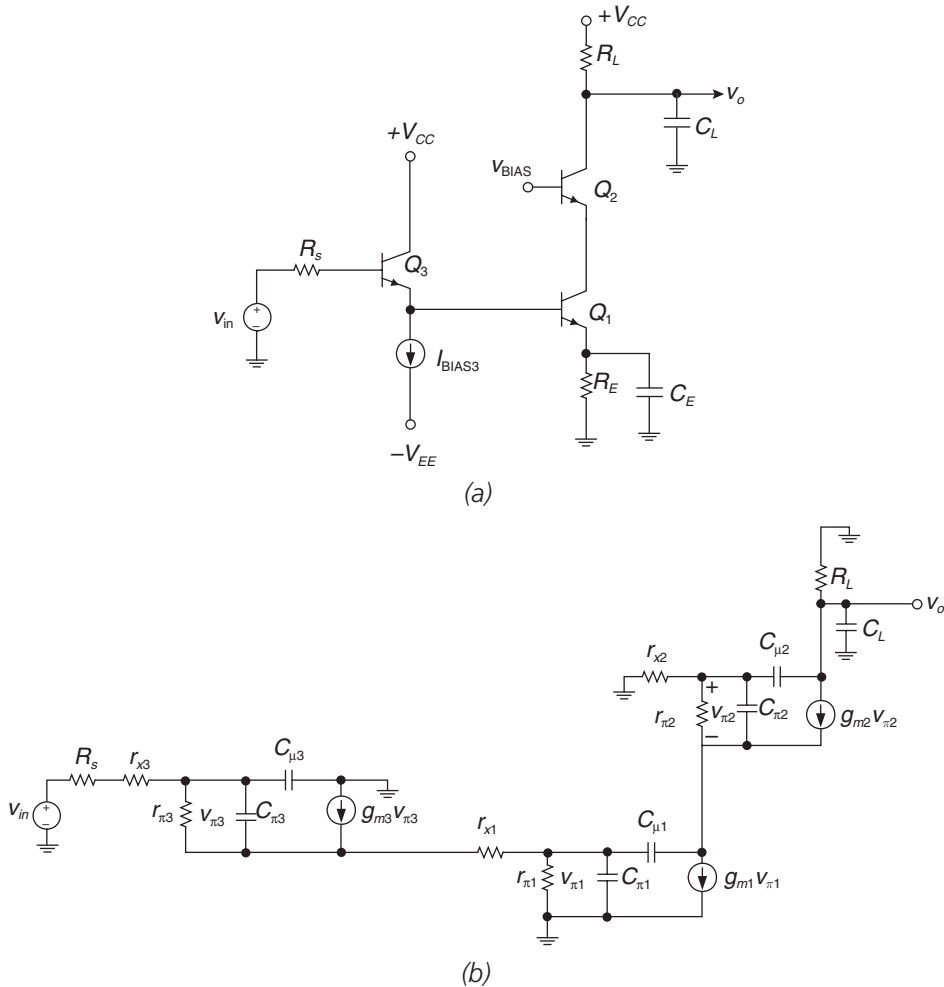


Figure 6-24: Third attempt (Try #3). (a) Circuit. (b) Small-signal model.

Table 6-4: OCTC summary for Try #1, #2 and #3.

	TRY #1	TRY #2	TRY #3
$\tau_{\pi 1}$	45 ns	45 ns	4.6 ns
$\tau_{\mu 1}$	442 ns	4.6 ns	0.5 ns
$\tau_{\pi 2}$	---	0.5 ns	0.5 ns
$\tau_{\mu 2}$	---	5 ns	3.2 ns
$\tau_{\pi 3}$	---	---	0.5 ns
$\tau_{\mu 3}$	---	---	4.2 ns
τ_{LOAD}	24 ns	24 ns	15 ns
$\sum \tau_{\text{oc}}$	511 ns	79 ns	28.5 ns
$\omega_{h, \text{est}}$	1.96 MR/s	12.7 MR/s	35.1 MR/s
$f_{h, \text{est}}$	311 kHz	2.0 MHz	5.6 MHz
f_h from PSPICE	315 kHz	2.5 MHz	9.4 MHz

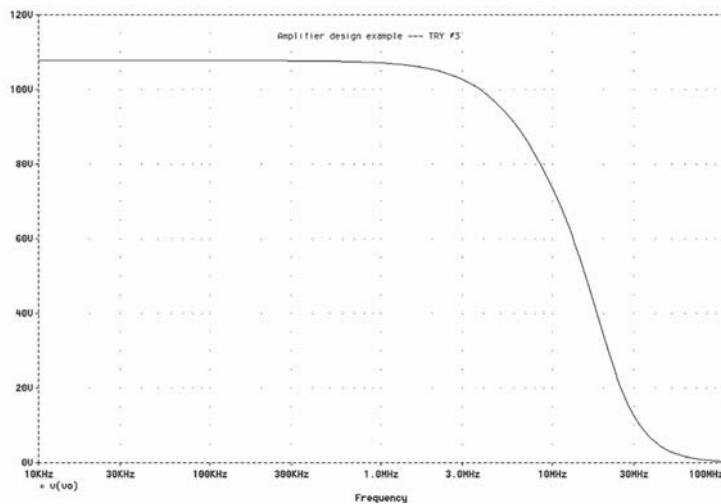
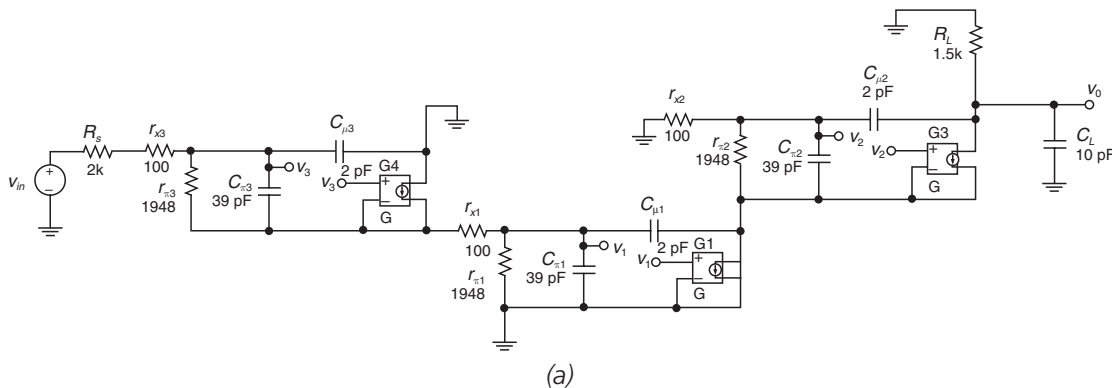


Figure 6-25: Frequency response for Try #3. (a) Circuit. (b) PSPICE frequency response result showing bandwidth of 9.4 MHz.

TRY #4—Emitter follower + common-emitter + cascode + emitter-follower

In Try #3, open-circuit time constants identified the interaction of the load capacitor with the load resistor R_L as the major bandwidth bottleneck. We can reduce the effects of C_L by adding a buffer emitter-follower Q_4 as shown in **Figure 6-26a** to buffer the load capacitor from the 1.5-k Ω load resistor. The addition of Q_4 doesn't affect the open-circuit time constants for Q_1 , Q_2 and Q_3 . We do need to calculate open-circuit time constants for Q_4 as well as a new OCTC for the load capacitor, using the small-signal model of **Figure 6-26b**.

For $C_{\pi 4}$, we find:

$$R_{\pi 4} \approx \frac{1}{g_{m4}} = 13\Omega \quad [6-68]$$

$$\tau_{\pi 4} = R_{\pi 4} C_{\pi 4} = (13)(39 \text{ pF}) = 0.5 \text{ ns}$$

Next, for $C_{\mu 4}$:

$$R_{\mu 4} = r_{x4} + R_L = 1600\Omega \quad [6-69]$$

$$\tau_{\mu 4} = R_{\mu 4} C_{\mu 4} = (1600)(2 \text{ pF}) = 3.2 \text{ ns}$$

For the load capacitor, we need to find the output resistance of the Q_4 emitter-follower:

$$r_{out,Q4} \approx \frac{R_L + r_{x4} + r_{\pi 4}}{1 + h_{fe4}} = \frac{1500 + 100 + 1948}{151} = 23.5\Omega \quad [6-70]$$

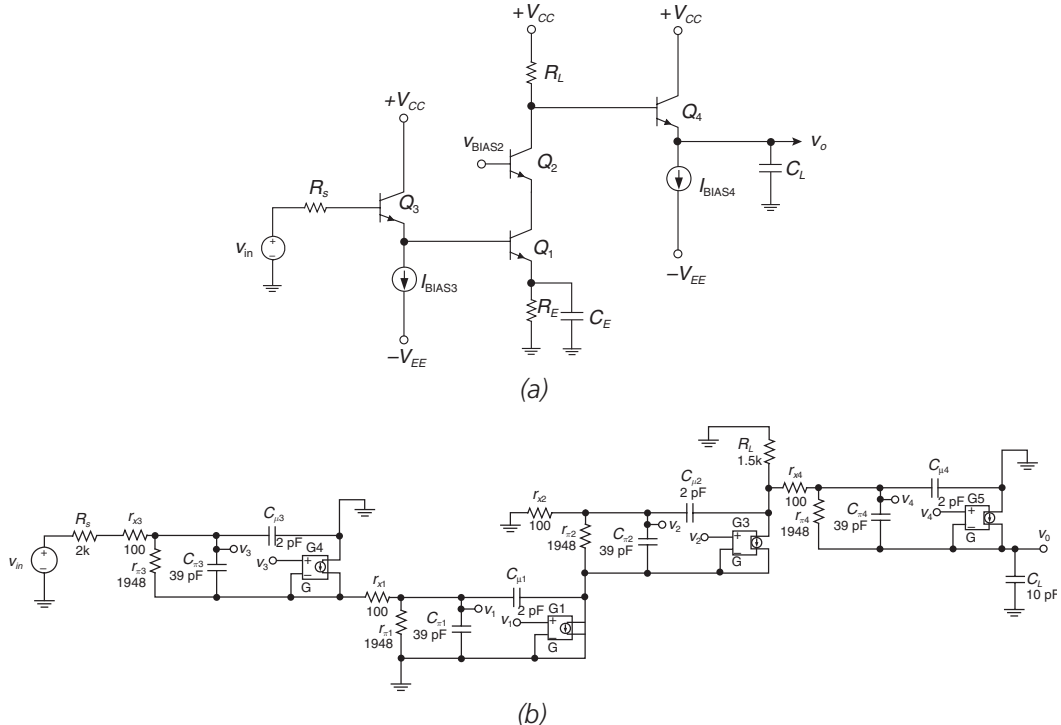


Figure 6-26: Fourth attempt (Try #4).

The resultant open-circuit time constant for the load capacitor is:

$$\tau_{LOAD} = r_{out,Q4} C_L = (23.5)(10 \text{ pF}) = 0.2 \text{ ns} \quad [6-71]$$

The sum of open-circuit time constants for this circuit is now 22 nanoseconds, and our estimate of bandwidth is:

$$\omega_h \approx \frac{1}{22 \text{ ns}} \approx 45.5 \text{ Mrad/sec}$$

$$f_h = \frac{\omega_h}{2\pi} \approx 7.23 \text{ MHz} \quad [6-72]$$

PSICE (**Figure 6-27**) shows that the gain is ~ -107.7 and the bandwidth is approximately 19.4 MHz. A summary of the open-circuit time constant results for Try #4 is shown in **Table 6-5**. We've finally met both the gain and bandwidth specifications for this amplifier.

Table 6-5: OCTC summary for Try #1, #2, #3 and #4.

	TRY #1	TRY #2	TRY #3	TRY#4
$\tau_{\pi1}$	45 ns	45 ns	4.6 ns	4.6 ns
$\tau_{\mu1}$	442 ns	4.6 ns	0.5 ns	0.5 ns
$\tau_{\pi2}$	---	0.5 ns	0.5 ns	0.5 ns
$\tau_{\mu2}$	---	5 ns	3.2 ns	3.2 ns
$\tau_{\pi3}$	---	---	0.5 ns	0.5 ns
$\tau_{\mu3}$	---	---	4.2 ns	4.2 ns
$\tau_{\pi4}$			---	0.5 ns
$\tau_{\mu4}$			---	3.2 ns
τ_{LOAD}	24 ns	24 ns	15 ns	0.2 ns
$\sum \tau_{oc}$	511 ns	79 ns	28.5 ns	22 ns
$\omega_{h,est}$	1.96 MR/s	12.7 MR/s	35.1 MR/s	45.4 MR/s
$f_{h,est}$	311 kHz	2.0 MHz	5.6 MHz	7.23 MHz
f_h from PSICE	315 kHz	2.5 MHz	9.4 MHz	19.4 MHz

We see that using the method of open-circuit time constants we are able to identify the major bandwidth bottlenecks and attack them. Furthermore, the OCTC method gives a pessimistic approximation of bandwidth, so that actual bandwidth will be higher. Note the following OCTC caveats:

- The bandwidth estimate is always conservative.
- The estimate is accurate if there is a dominant pole.
- The OCTC estimate can be over-pessimistic if there are dominant complex poles, since the method assumes real-axis poles.
- Not all capacitors are used in OCTC analysis (e.g., coupling and bypass capacitors).
 - OCTC applies only to high-frequency models.

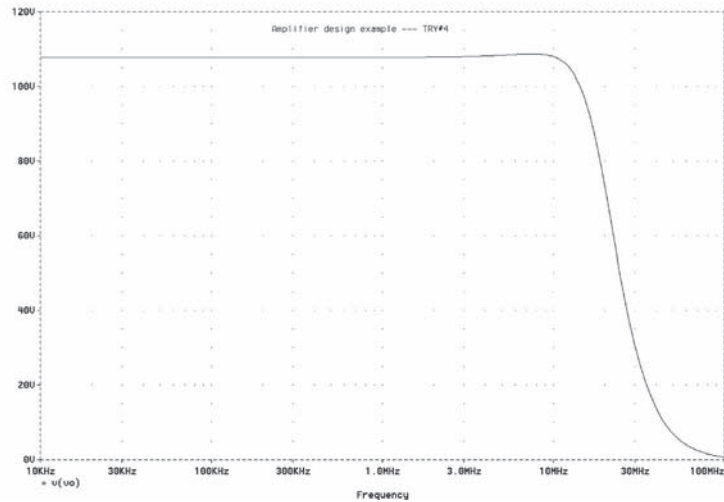


Figure 6-27: Frequency response of Try #4 showing bandwidth of 19.4 MHz.

- Each OCTC does *not* equate with a single system pole.
 - Remember, *sum* of pole time constants equals the *sum* of OCTCs, but the individual terms aren't necessarily equal.
- System zeros are not accounted for.

Chapter 6 Problems

Problem 6.1

Using the method of open-circuit time constants, estimate the bandwidth (in Hz) of the RC ladder in **Figure 6-28**.

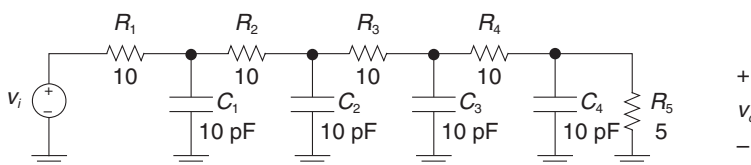


Figure 6-28: RC ladder for Problem 6.1.

Problem 6.2

For the circuit in **Figure 6-29**:

- Estimate the -3dB bandwidth using the method of open-circuit time constants.
- Find the -3dB bandwidth exactly by calculation or by using SPICE.

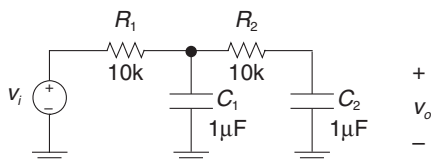
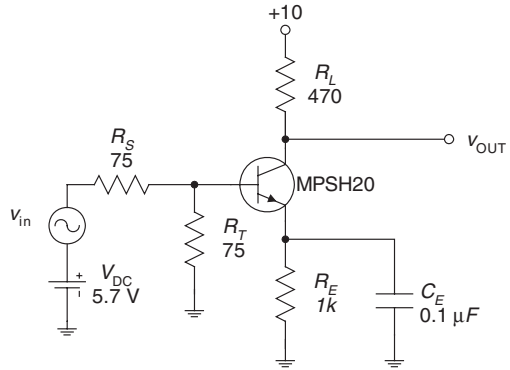


Figure 6-29: Circuit for Problem 6.2.

Problem 6.3

The transistor amplifier shown (**Figure 6-30**) is a high-bandwidth, common-emitter configuration used to amplify a video signal. A small-signal video input drives from a 75Ω source impedance and is terminated with a 75Ω termination resistor. The transistor is biased with a DC power supply $V_{\text{DC}} = 5.7$ volts. For the MPSH20 transistor you may assume $r_x = 20\Omega$, AC beta $h_{fe} \approx$ DC beta $h_{FE} = 25$, $C_\mu = 0.9$ pF and $f_T = 630$ MHz. Make reasonable approximations, state them, and justify them. Assume room temperature is 25°C and hence $kT/q = 26$ millivolts.

Figure 6-30: Common-emitter amplifier for Problem 6.3 and 6.4.



- Find the DC bias point values of collector current I_C and collector-base voltage V_{CB} . Find the power dissipated in the transistor.
- Calculate small-signal parameters g_m , r_π , C_π and C_μ , and draw the small-signal model. Include the emitter bypass capacitor C_E in your model.
- Now, assume that the operating frequency is high enough so that C_E acts as a short circuit, but low enough so that C_π and C_μ are open circuits. Redraw the small-signal model given these conditions, and calculate the midband voltage gain A_v .
- Using the method of open-circuit time constants, estimate the approximate high-frequency bandwidth f_H , in hertz.

Problem 6.4

When you build the circuit of Problem 6.3 and attach an oscilloscope to the output, you find that the measured bandwidth f_H is significantly lower than you expected based on your OCTC analysis. This effect is especially pronounced, since you know that the method of OCTC gives answers that are conservative. You trace down two causes:

- There are 2 picofarads of parasitic capacitance across the collector-base junction, due to the way you wired the circuit on the protoboard;
 - There are 10 picofarads of capacitance from v_{out} to ground, due to the loading effect of the scope probe.
- Given these two parasitic effects, what bandwidth f_H does the method of open-circuit time constants predict for your circuit?
 - In order to improve the bandwidth of your circuit, you fix the parasitic capacitance problem at the base-collector junction by directly soldering the collector lead to the load resistor without plugging the node into the protoboard. You may assume that this reduces the 2 pF of parasitic capacitance to zero at Q_1 's base-collector node. (Remember that you still have $C_{\mu 1}$ across this junction.) You also add an emitter-follower to isolate the load resistor from the output capacitive load due to the scope (**Figure 6-31**). Assume that the C_π time constant for Q_2 is 0.25 nanoseconds and the C_μ time constant for Q_2 is 0.4 nanoseconds. What value of f_H does open-circuit time constants predict?

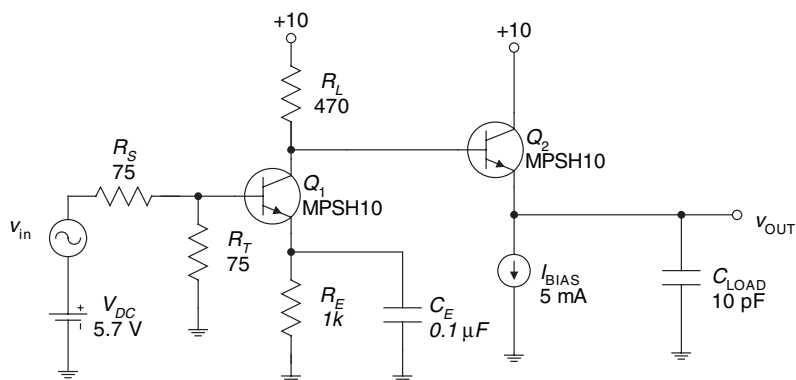


Figure 6-31: Common-emitter amplifier with output emitter-follower (Problem 6.4).

Problem 6.5

As we have seen, there is a relationship between bandwidth and risetime. The method of open-circuit time constants, since it enables us to estimate bandwidth, may also allow us to estimate the risetime of a single system, or of multiple systems in cascade. Consider a cascade of N amplifiers, each with a single pole with time constant τ .

- Using open-circuit time constants, estimate the bandwidth of the overall amplifier. From this bandwidth, estimate the risetime of the overall system.
- Using the risetime addition rule for systems in cascade,¹¹ i.e.:

$$\tau_{R, total} \approx \sqrt{\tau_{R1}^2 + \tau_{R2}^2 + \cdots + \tau_{RN}^2}$$

estimate the risetime. Comment on how this compares with the OCTC-derived estimate.

- Using SPICE for the case where $N = 5$ and $\tau = 1$, find the actual bandwidth and risetime. Compare the results to your previous calculations. Comment on the accuracy of the OCTC estimate.
- Now, repeat your calculations for the case with five systems in cascade, with four pole time constants $\tau = 1$, and one dominant time constant, $\tau = 10$. Comment on the accuracy of OCTC estimate for risetime.

¹¹ Note that this expression is not exact.

Problem 6.6

- (a) Assume that the circuit in **Figure 6-32** is driven by an AC voltage source. Using an intuitive approach, thought experiments, etc. sketch the shape of the Bode magnitude response of v_{out}/v_{in} .
- (b) Using the method of open-circuit time constants, find the high-frequency breakpoint f_H . Fill in f_L and f_H in your Bode plot.



Figure 6-32: Circuit for Problem 6.6.

References

- Davis, A., and Moustakas, E., "Analysis of active RC networks by decomposition," *IEEE Transactions on Circuits and Systems*, vol. 27, no. 5, May 1980, pp. 417–419.
- Fox, R. M., and Lee, S. G., "Extension of the open-circuit time-constant method to allow for transcapacitances," *IEEE Transactions on Circuits and Systems*, vol. 37, no. 9, September 1990, pp. 1167–1171.
- Gray, Paul E., and Searle, Campbell L., *Electronic Principles Physics, Models and Circuits*, John Wiley, 1969.
- Gray, Paul R., and Meyer, Robert G., *Analysis and Design of Analog Integrated Circuits*, 2d edition, John Wiley, 1984.
- Hunt, F.V., and Hickman, R.W., "On Electronic Voltage Stabilizers," *Review of Scientific Instruments*, vol. 10, Jan. 1939, pp. 6–21.
- Rathore, T. S., "Generalized Miller theorem and its applications," *IEEE Transactions on Education*, vol. 32, no. 3, August 1989, pp. 386–390.
- Thompson, Marc T., "Design Linear Circuits Using OCTC Calculations," *Electronic Design (Special Analog Issue)*, June 24, 1993, pp. 41–47.
- , "Network Tricks Aid in OCTC," *Electronic Design*, December 16, 1993, pp. 67–70.
- Thornton, R. D., Searle, C. L., Pederson, D. O., Adler, R. B., and Angelo, E. J., Jr., *Multistage Transistor Circuits* (SEEC Vol. 5), John Wiley, 1965.

Advanced Transistor Amplifier Techniques

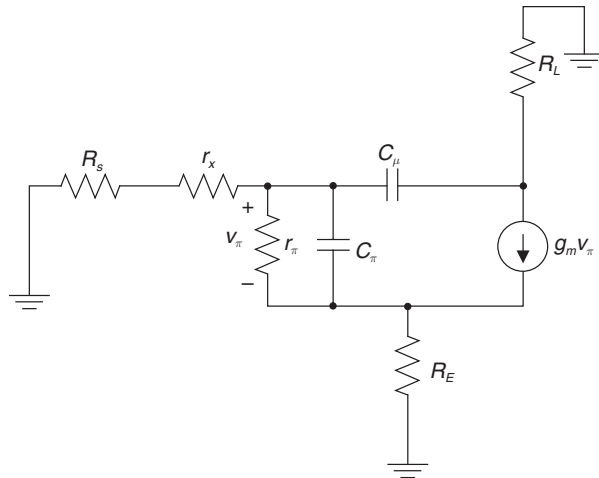
In This Chapter

- Various and sundry issues related to advanced amplifier design techniques are discussed.

Worst-Case Open-Circuit Time Constant Calculations

The most complicated open-circuit time constant calculation is found for a transistor amplifier that has a source resistance (R_s), emitter resistance (R_E) and collector load resistance (R_L), as shown in **Figure 7-1**. In the following section, we'll figure out the open-circuit resistances for C_π and C_μ for this worst-case circuit.

Figure 7-1: Worst-case open-circuit time constant circuit with source resistance R_s , emitter resistor R_E and collector load resistor R_L .



The worst-case open-circuit time constants circuit for C_π is shown in **Figure 7-2a**. The mathematics rapidly gets difficult if we blindly begin writing node equations. First, let's recognize that the test voltage source v_i sets the value of v_π . Since the v_i generator sets the value of the dependent¹ current source, we *can* use superposition to find the test current.

¹ Normally, we can't use superposition in circuits with dependent sources. However, in this special case we can, since the test voltage source v_i directly sets the dependent current source value as well.

Using the circuit of **Figure 7-2b**, we'll find the test current due to the test voltage source, with the $g_m v_t$ dependent generator set to zero as an open circuit. By inspection, we see that the resistance across the v_t test source is r_π in parallel with $R_s + r_x + R_E$. Therefore, the test current i_{t1} due to the v_t generator only is:

$$i_{t1} = \frac{v_t}{r_\pi} + \frac{v_t}{R_s + r_x + R_E} \quad [7-1]$$

To find the current due to the $g_m v_t$ test generator only, we use the circuit of **Figure 7-2c**. Using this circuit, and recognizing the current divider we find:

$$i_{t2} = g_m v_t \frac{R_E}{R_s + r_x + R_E} = v_t \frac{g_m R_E}{R_s + r_x + R_E} \quad [7-2]$$

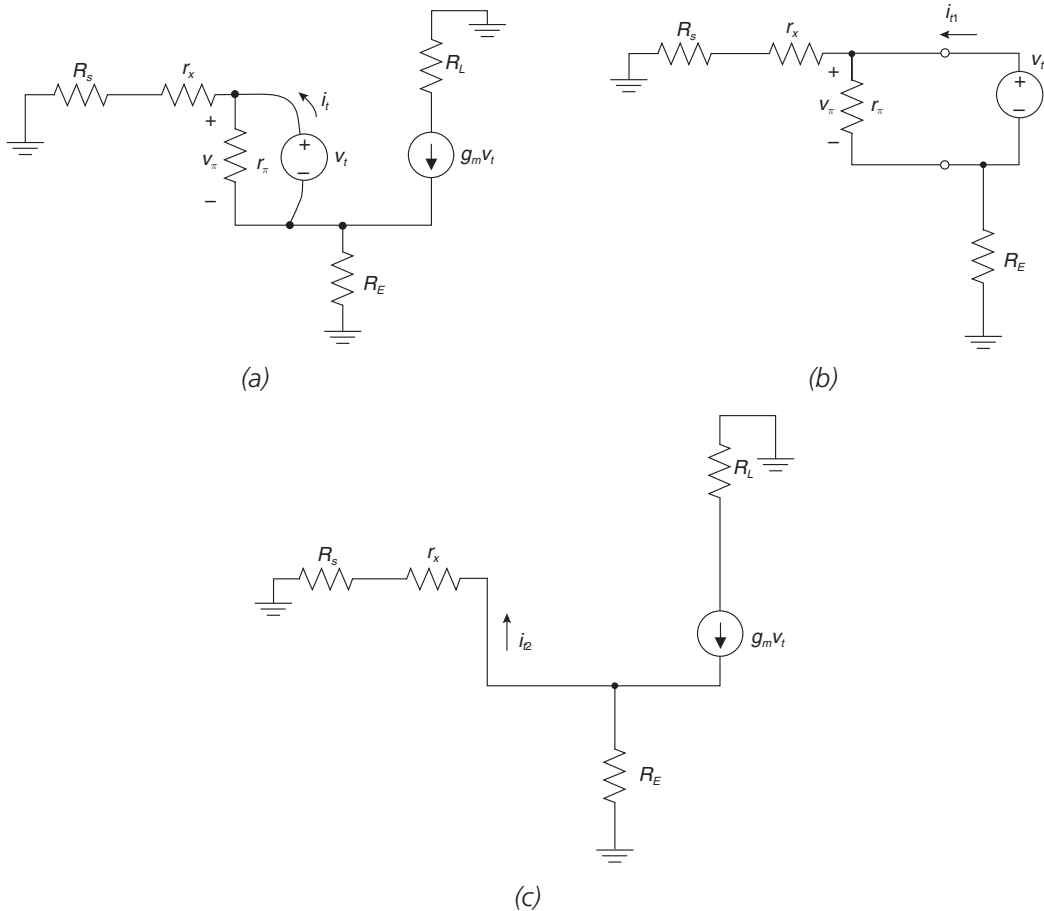


Figure 7-2: Circuits for finding worst-case OCTC for C_π . (a) Original circuit. (b) Circuit for finding contribution due to test voltage source v_t only. (c) Circuit for finding contribution due to $g_m v_t$ generator only.

The total test current i_t is the sum of i_{t1} and i_{t2} , and is:

$$i_t = i_{t1} + i_{t2} = v_t \left(\frac{1}{r_\pi} + \frac{1 + g_m R_E}{R_s + r_x + R_E} \right) \quad [7-3]$$

We find that the open-circuit resistance for C_π is:

$$R_{o\pi} = \frac{v_t}{i_t} = r_\pi \left\| \left(\frac{R_s + r_x + R_E}{1 + g_m R_E} \right) \right. \quad [7-4]$$

Let's note some important limiting cases. In the case of a large emitter resistor with $R_E \gg r_x$ and R_s , we find:

$$R_{o\pi} \approx \frac{1}{g_m} \text{ if } R_E \gg R_s, r_x \quad [7-5]$$

In the case of a small emitter resistor² with $R_E \ll r_x$ and R_s and with $g_m R_E \ll 1$, we find:

$$R_{o\pi} \approx r_\pi \left\| (R_s + r_x) \right. \quad [7-6]$$

The worst-case open-circuit time constant circuit for C_μ (**Figure 7-3a**) initially looks like a difficult problem as well. However, let's transform the circuit to a simpler form (**Figure 7-3b**). The circuit of **Figure 7-3b** is equivalent to the circuit of **Figure 7-3a** provided we choose the correct values³ of R_1 and G_M . The value of R_1 is easy to find; it's just the input resistance looking into the top of r_π , so we find:

$$R_1 = r_\pi + (1 + h_{fe}) R_E \quad [7-7]$$

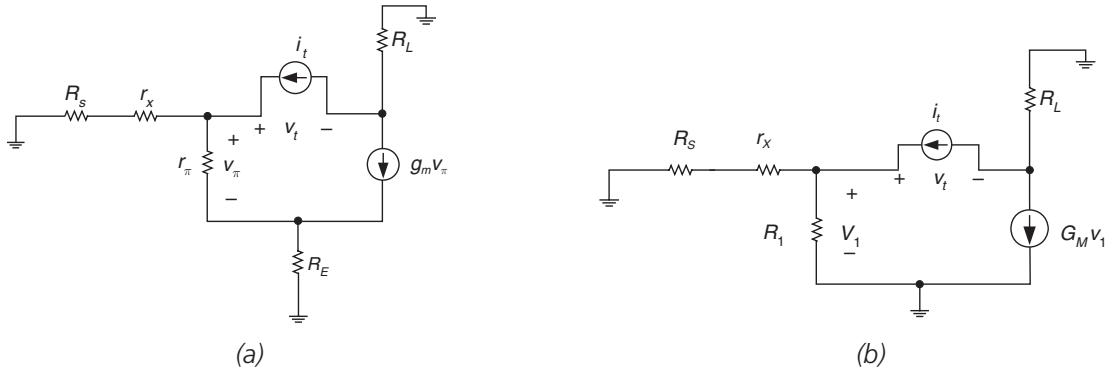


Figure 7-3: Circuit for finding worst-case OCTC for C_μ . (a) Original circuit. (b) Circuit redrawn in equivalent form that is easier to solve.

² We worked out this case in the previous chapter, for the common-emitter amplifier with $R_E = 0$.

³ A similar development is given in Gray, Hurst, Lewis and Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th edition, pp. 197–200.

Chapter 7

Next, we need to find the value of the new dependent generator. We can make use of the fact that the output currents from the old and transformed circuits are the same, or:

$$g_m v_\pi = G_M v_1 \quad [7-8]$$

Let's imagine using a test current source i_t to solve for v_π in **Figure 7-3a**, and for v_1 in **Figure 7-3b**. We find⁴ for v_π :

$$v_\pi = i_t \left(\frac{R_s + r_x}{R_s + r_x + r_\pi + (1 + h_{fe})R_E} \right) r_\pi \quad [7-9]$$

Next, we solve for v_1 :

$$v_1 = i_t \left(\frac{(R_s + r_x)R_1}{R_s + r_x + R_1} \right) = \frac{(R_s + r_x)(r_\pi + (1 + h_{fe})R_E)}{R_s + r_x + r_\pi + (1 + h_{fe})R_E} \quad [7-10]$$

Next, we solve⁵ for G_M :

$$G_M = \frac{g_m v_\pi}{v_1} = g_m \frac{\left(\frac{(R_s + r_x)}{R_s + r_x + r_\pi + (1 + h_{fe})R_E} \right) r_\pi}{\left(\frac{(R_s + r_x)(r_\pi + (1 + h_{fe})R_E)}{R_s + r_x + r_\pi + (1 + h_{fe})R_E} \right)} = \frac{g_m}{1 + \frac{R_E}{r_\pi} + g_m R_E} \approx \frac{g_m}{1 + g_m R_E} \quad [7-11]$$

To summarize, the worst-case open-circuit resistance for C_μ is:

$$R_{o\mu} = R_{EQ1} + R_L + G_M R_{EQ1} R_L \quad [7-12]$$

$$R_{EQ1} = (R_s + r_x) \parallel R_1 = (R_s + r_x) \parallel (r_\pi + (1 + h_{fe})R_E)$$

$$G_M \approx \frac{g_m}{1 + g_m R_E}$$

⁴ This is done by calculating a current divider. The fraction of the it current that travels through r_π is:

$$\left(\frac{R_s + r_x}{R_s + r_x + r_\pi + (1 + h_{fe})R_E} \right)$$

⁵ Yes, I know it's a mess, but follow it through. It arrives at a nice simple result. Note that $g_m \gg 1/r_\pi$.

Example 7.1: Estimating bandwidth of common-emitter amplifier with emitter degeneration

Shown in **Figure 7-4** is a common-emitter amplifier with *emitter degeneration*. Note that there's a finite resistor R_E in the emitter leg of the transistor. Emitter degeneration lowers the gain of the common-emitter amplifier, but extends the bandwidth by partially bootstrapping the base-emitter capacitance and by lowering the C_μ time constant by lowering the gain. We can use the previously derived worst-case open-circuit resistance calculations to estimate the bandwidth of this amplifier.

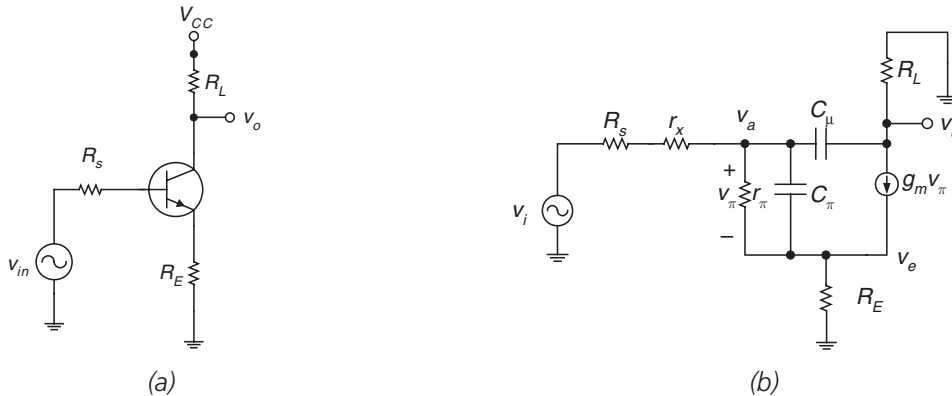
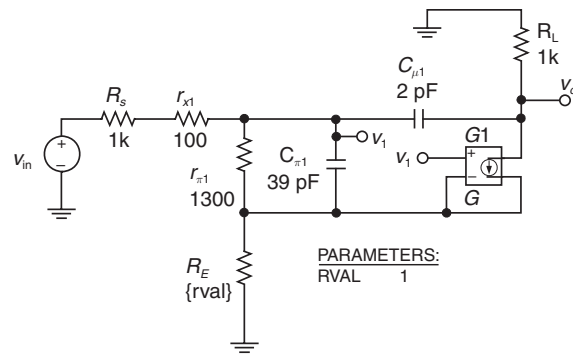


Figure 7-4: Common-emitter amplifier with emitter degeneration, biasing details omitted. (a) Circuit. (b) High-frequency incremental model.

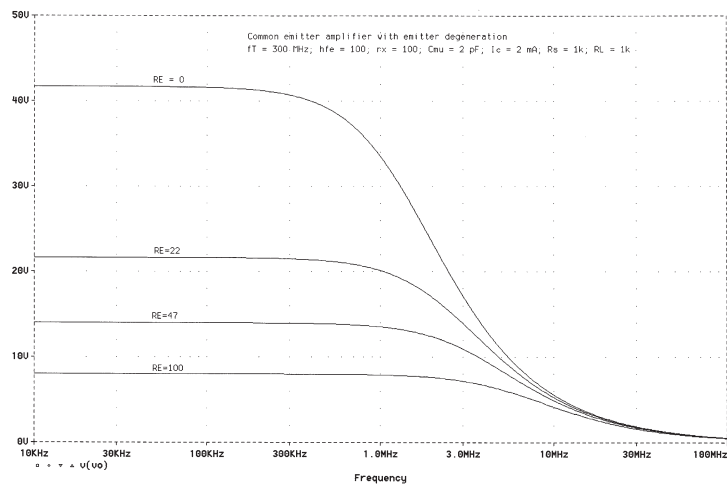
Assume that the transistor is biased at a collector current $I_C = 2$ milliamps, and that the transistor has small-signal current gain $h_{fe} = 100$, base resistance $r_x = 100\Omega$ and collector-base capacitance $C_\mu = 2$ picofarads. The load resistor is $R_L = 1000\Omega$ and the emitter resistor R_E is variable. The gain and open-circuit time constants estimate of bandwidth for this circuit is shown in **Table 7-1**. Note that as the emitter resistor is increased, the gain decreases while bandwidth increases. This gives the designer another degree-of-freedom in setting gain and bandwidth in common-emitter amplifiers. These results are confirmed by a PSPICE simulation (**Figure 7-5**).

Table 7-1: Common-emitter amplifier with emitter degeneration results. The computation $f_{h,est}$ is the estimated -3dB bandwidth using open-circuit time constants.

R_E	Gain (calculated)	OCTC for C_π (ns)	OCTC for C_μ (ns)	$\omega_{h,est}$ (Mr/sec)	$f_{h,est}$ (MHz)	-3dB from PSPICE (MHz)
0	-41.7	23.1	94.9	8.48	1.35	1.35
22 Ω	-21.6	12.2	51.6	15.7	2.49	2.52
47 Ω	-14.0	8.1	34.9	23.3	3.7	3.75
100 Ω	-8.0	4.8	21.8	37.6	6.0	6.10



(a)



(b)

Figure 7-5: Common-emitter amplifier with emitter degeneration.

(a) PSPICE circuit. Note that the value of R_E is variable.

(b) Circuit simulations for $R_E = 0, 22\Omega, 47\Omega$ and 100Ω .

Example 7.2: Gain of differential amplifier with emitter degeneration

The model of the common-emitter amplifier with emitter degeneration can be used to analyze the gain of the differential amplifier with emitter degeneration (**Figure 7-6a**) when operated in the differential mode. By half-circuit techniques, we see that in the differential mode,⁶ the circuit is equivalent to the common-emitter amplifier with emitter degeneration (**Figure 7-6b**). Inclusion of emitter resistors R_E increases the range of differential input voltages over which the differential amplifier provides approximately linear gain.⁷

Shown in **Figure 7-7** is a PSPICE simulation of a differential amplifier with emitter degeneration. The base of transistor Q_2 is grounded and the voltage at the base of transistor Q_1 is swept from -500 millivolts to $+500$ millivolts, and the voltage at the base of Q_2 is swept simultaneously from $+500$ millivolts to -500 millivolts. Results are shown in the simulation of **Figure 7-7b**. With $R_E = 0$, the amplifier is a standard differential amplifier, and the output swings over its full range when the input varies a few kT/q . With $R_E = 100\Omega$ and 220Ω , the amplifier has linear gain over a much wider input voltage range.

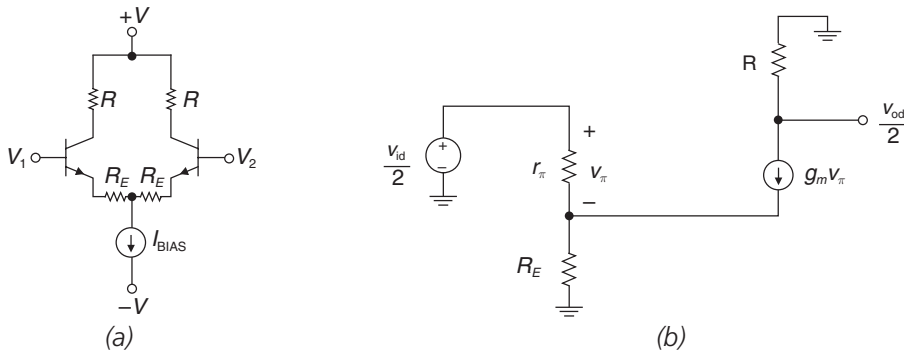
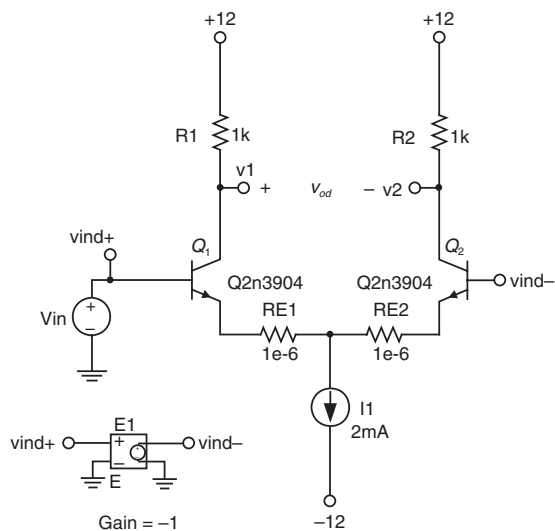


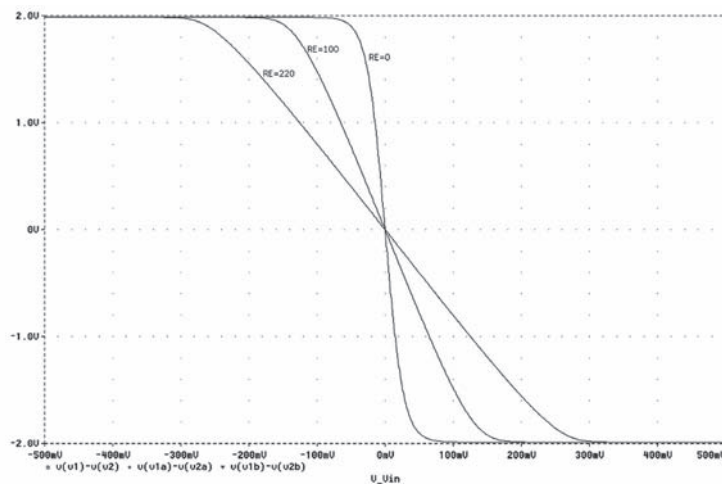
Figure 7-6: Differential amplifier with emitter degeneration. (a) Circuit. (b) Low-frequency small-signal model for the differential mode of operation.

⁶ In the differential mode, the input to one of the transistors is $v_{id}/2$ and the output of one of the transistors is $v_{od}/2$.

⁷ However, there are tradeoffs—an increase in noise due to the resistors, as well as a lowering of differential gain and CMRR. A full discussion is given in Grebene's *Bipolar and MOS Analog Integrated Circuit Design*, pp. 221–224.



(a)



(b)

Figure 7-7: Differential amplifier with emitter degeneration. (a) PSPICE circuit. (b) Simulation result with $R_E = 0, 100\Omega$ and 220Ω and input v_{in} swept from -500 mV to $+500$ mV.

High-Frequency Output and Input Impedance of Emitter-Follower Buffers

An ideal buffer has very high input impedance and very low output impedance. The output and input impedance of an emitter-follower varies with frequency due to the effects of C_π and C_μ . Using the circuits of **Figure 7-8**, we'll further investigate the high-frequency output

impedance of the emitter-follower. We found in Chapter 5, using the circuit of **Figure 7-8b**, that the low-frequency output resistance of the emitter-follower is:

$$Z_{out}|_{\omega \rightarrow 0} = \left(\frac{R_s + r_x + r_\pi}{1 + h_{fe}} \right) \approx \left(\frac{R_s + r_x}{h_{fe}} \right) + \frac{1}{g_m} \quad [7-13]$$

We saw earlier that the small-signal current gain of the transistor decreases at very high frequencies. At very high frequencies, C_π shorts out r_π and the $g_m v_\pi$ generator is turned off as well, resulting in the incremental model of **Figure 7-8c**. Ignoring the effects of C_μ , the output impedance at very high frequencies is:

$$Z_{out}|_{\omega \rightarrow \infty} = R_s + r_x \quad [7-14]$$

If we wanted, we could find the output impedance in closed form by using the following methodology. We can reuse the previous result for the low-frequency output impedance of the emitter-follower. However, if we replace r_π in the original equation with an impedance representing r_π in parallel with C_π , we can solve for the output impedance as a function of frequency:

$$Z_{out}(s) = \left(\frac{R_s + r_x + z_\pi(s)}{1 + h_{fe}} \right) \quad [7-15]$$

$$z_\pi(s) = \frac{r_\pi C_\pi}{r_\pi C_\pi s + 1}$$

However, this gets mathematically complicated, so let's resort to an approximate result, which we can find by inspection with a minimum of calculation.

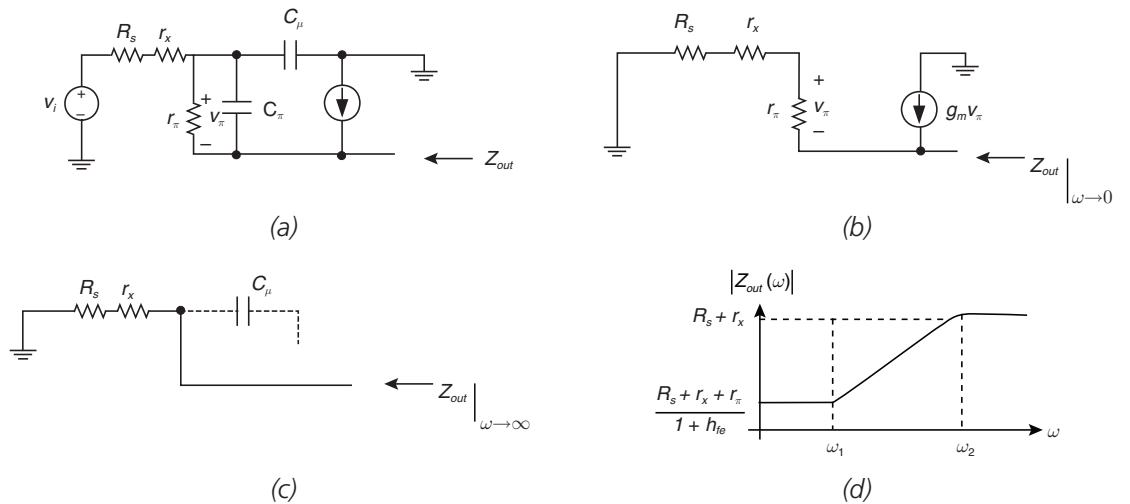


Figure 7-8: Analysis of high-frequency output impedance of emitter-follower. (a) Incremental circuit. (b) Low-frequency model. (c) High-frequency model. (d) Impedance plot of output impedance showing that output impedance increases over a finite frequency range from ω_1 to ω_2 . In this region, the output impedance of the emitter-follower is inductive.

A plot of emitter-follower output impedance magnitude is shown in **Figure 7-8d**, assuming⁸ that collector current is high enough so that $1/g_m < R_s + r_x$. We note that between frequencies ω_1 and ω_2 the output impedance increases linearly with frequency, and hence is inductive in this frequency range.

The frequency ω_1 when the impedance begins to increase is the frequency at which C_π begins shorting out r_π . This frequency is approximately:

$$\omega_1 \approx \frac{1}{r_\pi C_\pi} \quad [7-16]$$

The frequency ω_2 is approximately h_{fe} times this frequency:

$$\omega_2 \approx h_{fe} \omega_1 \approx \frac{h_{fe}}{r_\pi C_\pi} \approx \frac{g_m}{C_\pi} \quad [7-17]$$

An equivalent circuit that mimics this frequency-dependent impedance is shown in **Figure 7-9**. We can fit the parameters R_1 and R_2 :

$$\begin{aligned} R_1 &= R_s + r_x \\ R_1 \parallel R_2 &= \frac{R_s + r_x + r_\pi}{1 + h_{fe}} \end{aligned} \quad [7-18]$$

In order to find the inductor value L we note that the time constants of the original transistor circuit and the inductive circuit model must be the same. The open-circuit resistance facing C_π is approximately $1/g_m$. The resistance facing the inductor L is $R_1 + R_2$. Hence we can find L by the following method of equating the RC and L/R time constants of the two circuits:

$$\frac{C_\pi}{g_m} = \frac{L}{R_1 + R_2} \Rightarrow L \approx \left(\frac{C_\pi}{g_m} \right) (R_1 + R_2) \quad [7-19]$$

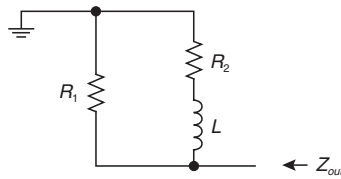


Figure 7-9: Model of output impedance $Z_{out}(s)$ of emitter-follower showing parameters L , R_1 and R_2 .

⁸ This assumption is that the high-frequency impedance is higher than the low-frequency impedance. A necessary condition for this to be true is:

$$\frac{R_s + r_x}{h_{fe}} + \frac{1}{g_m} < R_s + r_x$$

This usually occurs at modest collector currents.

Example 7.3: Emitter-follower output impedance

Let's do a numerical example. Assume we have an emitter-follower with $R_s = 1\text{ k}\Omega$, $r_x = 200\Omega$, $I_c = 2\text{ milliamps}$, $h_{fe} = 150$, $g_m = 0.077\Omega^{-1}$, $r_\pi = 1950\Omega$, $C_\mu = 2\text{ picofarads}$, $f_T = 300\text{ MHz}$ and $C_\pi = 39\text{ picofarads}$. The incremental model for this emitter-follower is shown in **Figure 7-10a**. Our inductive model predicts:

$$\begin{aligned} R_1 &= R_s + r_x = 1200\Omega \\ R_1 \parallel R_2 &= \frac{R_s + r_x + r_\pi}{1 + h_{fe}} = 20.9\Omega \Rightarrow R_2 \approx 20.9\Omega \\ L &\approx \left(\frac{C_\pi}{g_m} \right) (R_1 + R_2) = \left(\frac{39\text{pf}}{0.077} \right) (1220) \approx 0.61\text{ }\mu\text{H} \end{aligned} \quad [7-20]$$

This model predicts an impedance that starts to rise at a frequency $\sim R_2/L$, or at 34.3 Mrad/sec (5.5 MHz).

We can simulate this with PSPICE by adding an AC test current source and measuring the resultant test voltage v_o , as in **Figure 7-10b**. The simulation result (**Figure 7-10c**) shows that the low-frequency impedance is approximately 20Ω as expected. However, the high-frequency impedance (at frequencies above approximately 100 MHz) doesn't reach our expected high-frequency limit of 1200Ω . This is due to the loading effect of C_μ , which our simplified analysis has ignored. At very high frequencies, C_μ causes the output impedance to once again decrease. So, in this design example, the models show that the range over which the output impedance is inductive is approximately 5 MHz to 100 MHz .

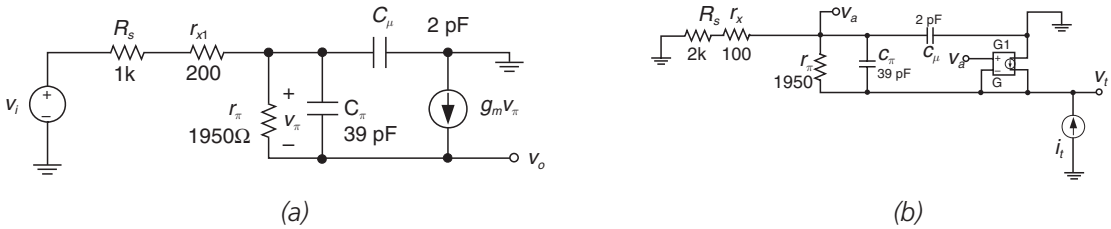
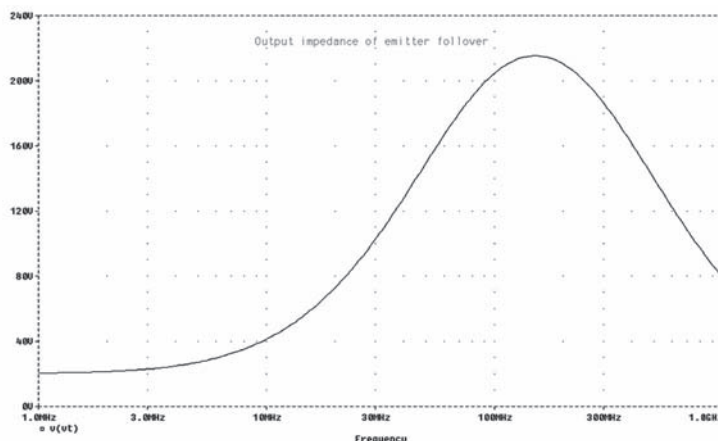


Figure 7-10: Analysis of high-frequency output impedance of emitter-follower using PSPICE. (a) High-frequency incremental circuit including r_x and C_μ . (b) PSPICE circuit, where an AC current source of value $i_t = 1$ drives the emitter so that we can find the emitter output impedance. (Continued on following page.)

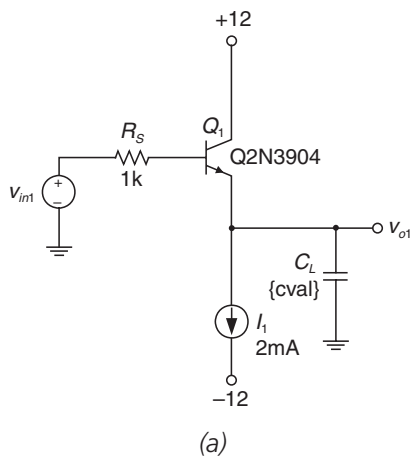


(c)

Figure 7-10 (continued): (c) PSPICE simulation, showing region of inductive output impedance, followed by impedance rolloff due to effects of C_μ .

One ramification of the inductive output impedance is that you can have gain peaking if you drive capacitive loads. Shown in **Figure 7-11a** is an emitter-follower driving a capacitive load.⁹ The PSPICE simulation (**Figure 7-11b**) shows gain peaking for modest capacitive loading. Note that the peak value occurs at frequencies where we expect the output of the emitter-follower to have inductive output impedance.¹⁰

Figure 7-11: Emitter-follower driving capacitive load. (a) Circuit using 2N3904 transistor. (Continued on following page.)



(a)

⁹ In this simulation, we use the 2N3904 transistor, which has parameters (f_T , h_{fe} , etc.) similar to that of the previous example.

¹⁰ We could damp this by adding a resistor in series with the load capacitor.

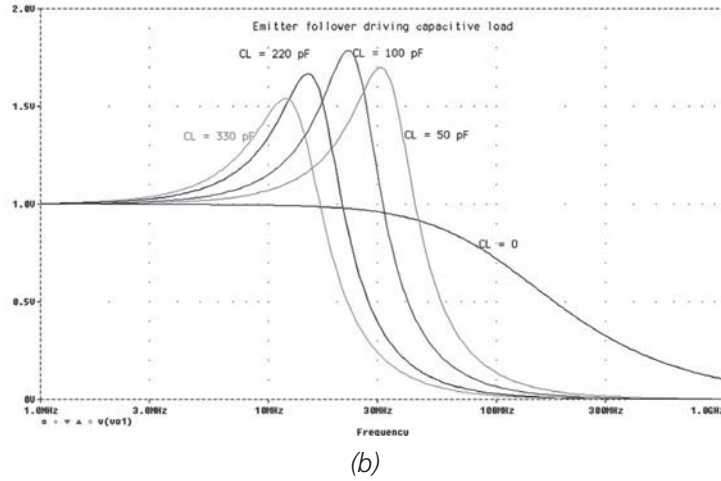


Figure 7-11 (continued): (b) PSPICE simulation of gain of this emitter-follower with load capacitor values $C_L = 0, 50 \text{ pF}, 100 \text{ pF}, 220 \text{ pF}$ and 330 pF .

Example 7.4: Emitter-follower input impedance

We'll next investigate the input impedance of the emitter-follower by using the simplified incremental model¹¹ of **Figure 7-12a**. Using a similar methodology as in Chapter 5 when we found the low-frequency input impedance of the emitter-follower, we find that the input impedance is:

$$Z_{in}(s) = \frac{v_t(s)}{i_t(s)} = z_\pi(s) + (1 + h_{fe}(s))R_E = z_\pi(s) + (1 + g_m z_\pi(s))R_E \quad [7-21]$$

$$z_\pi(s) = \frac{r_\pi}{r_\pi C_\pi s + 1}$$

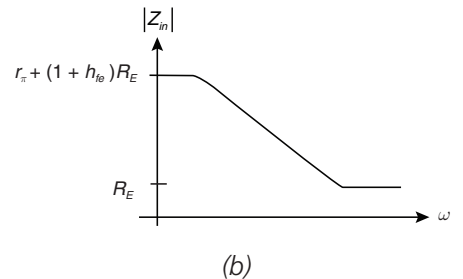
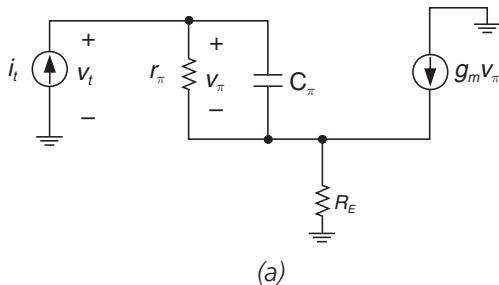


Figure 7-12: Simplified model for finding input impedance $Z_{in}(s)$ of emitter-follower. (a) Original circuit assuming $C_\mu = 0$ and $r_x = 0$. (b) Plot of magnitude of input impedance.

¹¹ It's simplified because we're ignoring r_x and C_μ throughout the following discussion. This makes the math much easier and gives us some insight into the physical operation of the circuit.

We find that at very low frequencies, the input impedance is:

$$Z_{in, \omega \rightarrow 0} = r_{\pi} + (1 + h_{fe})R_E \quad [7-22]$$

At very high frequencies, C_{π} shorts out r_{π} and disables the dependent current source, resulting in:

$$Z_{in, \omega \rightarrow \infty} = R_E \quad [7-23]$$

A plot of the magnitude of the input impedance of the emitter-follower as a function of frequency is shown in **Figure 7-12b**. Note that at very high frequencies, the impedance $Z_{in}(\omega)$ will further decrease due to the shunting effect of C_{μ} , which we ignored in the previous analysis.

Next, let's see what happens when we capacitively load the emitter-follower as in **Figure 7-13**. The next logical progression is:

$$\begin{aligned} Z_{in}(s) &= z_{\pi}(s) + (1 + g_m z_{\pi}(s))Z_E(s) \\ z_{\pi}(s) &= \frac{r_{\pi}}{r_{\pi} C_{\pi} s + 1} \\ Z_E(s) &= \frac{R_E}{R_E C_L s + 1} \end{aligned} \quad [7-24]$$

Following through with the algebra, we find:

$$\begin{aligned} Z_{in}(s) &= \frac{r_{\pi}}{r_{\pi} C_{\pi} s + 1} + \left(1 + g_m \left(\frac{r_{\pi}}{r_{\pi} C_{\pi} s + 1} \right) \right) \left(\frac{R_E}{R_E C_L s + 1} \right) \\ &= \left(r_{\pi} + (1 + h_{fe})R_E \right) \left(\frac{\frac{r_{\pi} R_E (C_L + C_{\pi}) s}{(r_{\pi} + (1 + h_{fe})R_E)} + 1}{(r_{\pi} C_{\pi} s + 1)(R_E C_L s + 1)} \right) \end{aligned} \quad [7-25]$$

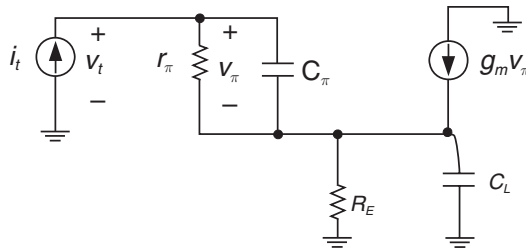


Figure 7-13: Incremental model of capacitively loaded emitter-follower for calculation of input impedance.

Normally, R_E is a large value and $h_{fe}R_E \gg r_\pi$, so we can approximate this input impedance as:

$$Z_{in}(s) \approx (r_\pi + (1 + h_{fe})R_E) \left(\frac{\frac{(C_L + C_\pi)s}{g_m} + 1}{(r_\pi C_\pi s + 1)(R_E C_L s + 1)} \right) \quad [7-26]$$

Now let's pause and look at this result. First, there's a term out front $(r_\pi + (1 + h_{fe})R_E)$ which is the low-frequency input impedance, as expected. Secondly, we find that there are two poles and a zero at frequencies:

$$\begin{aligned} \omega_{p1} &= -\frac{1}{R_E C_L} \\ \omega_{p2} &= -\frac{1}{r_\pi C_\pi} \\ \omega_z &= -\frac{g_m}{C_L + C_\pi} \end{aligned} \quad [7-27]$$

Let's plot the pole-zero plot and Bode plot for two different cases. First, for relatively small C_L , with $C_L < h_{fe}C_\pi$ (**Figure 7-14a**), we find that the zero frequency ω_z is at a frequency higher than ω_{p2} . This means that the angle of the input impedance dips below -90 degrees for some frequency range. This means that in this frequency range, the real part of the input impedance is negative.¹² In other words, in this frequency range the negative real input impedance can help sustain an oscillation.

Secondly, for relatively large C_L (**Figure 7-14b**) we find that the zero frequency ω_z is at a frequency higher than ω_{p2} . Thus, the real part of the input impedance never becomes negative.

This result shows that emitter-followers can exhibit strange behavior—circuits containing emitter-followers can have gain peaking, or in some cases can even oscillate.¹³ One design strategy that is sometimes taken is to ensure that the base of an emitter-follower has some finite resistance in series. As seen in **Figure 7-15**, we've put 22Ω in series¹⁴ with the base of the emitter-follower.

¹² A negative resistance can source power, while a positive resistance can only dissipate power.

¹³ A negative resistor delivers power to an external circuit. A positive resistor can only dissipate power. If the negative resistor delivers power to an LC circuit, an oscillation can result.

¹⁴ This value is a good starting point. Without any base resistance, emitter-followers built with 2N3904s, 2N2222s, 2N3906s, etc. or other signal transistors can oscillate at very high frequencies (~ 100 MHz or higher). We could also use a lossy ferrite bead in series with the base.

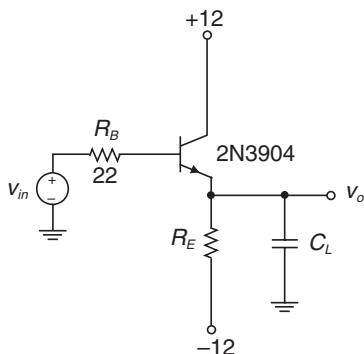
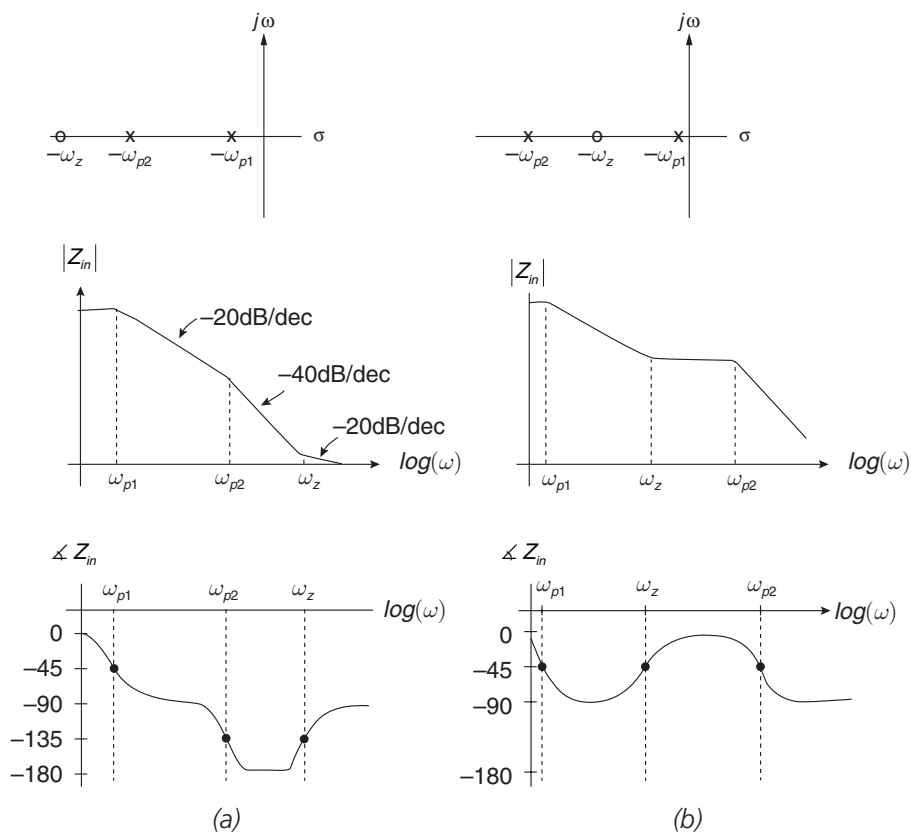


Figure 7-15: Resistor R_B added to base of emitter-follower to reduce chances of oscillation.

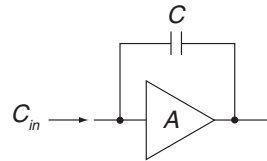
Bootstrapping

We have seen earlier in our discussion of the Miller effect that the effects of capacitance can be multiplied if the capacitance is wrapped around a negative gain. Referring to **Figure 7-16**, remember our result from a previous chapter that the capacitance looking into the input terminal of this amplifier is:

$$C_{in} = C(1 - A) \quad [7-28]$$

This result suggests another way to reduce the effects of circuit capacitances. What happens if the amplifier has a gain of +1? With a gain of +1, the capacitance looking into the terminals of the amplifier is exactly zero. This methodology can be used to reduce the effects of capacitances and hence to enhance bandwidth.

Figure 7-16: Circuit illustrating Miller effect.



Example 7.5: Bootstrapping an emitter-follower

We can illustrate bootstrapping by looking at an emitter-follower circuit (**Figure 7-17**).

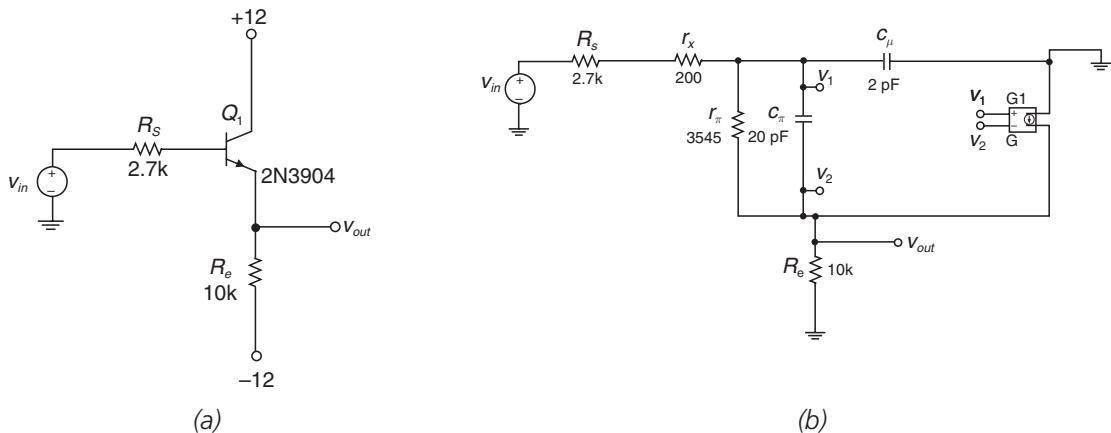


Figure 7-17: Emitter-follower for bootstrapping example.
(a) Circuit. (b) Small-signal model used for PSPICE simulation.

A detailed open-circuit time constant calculation shows that the open-circuit time constant for C_{π} is approximately 0.6 nanoseconds, and the open-circuit time constant for C_{μ} is approximately 6 nanoseconds. An estimate of bandwidth is therefore 25 MHz. A PSPICE simulation (**Figure 7-18**) shows a bandwidth of ~27 MHz.

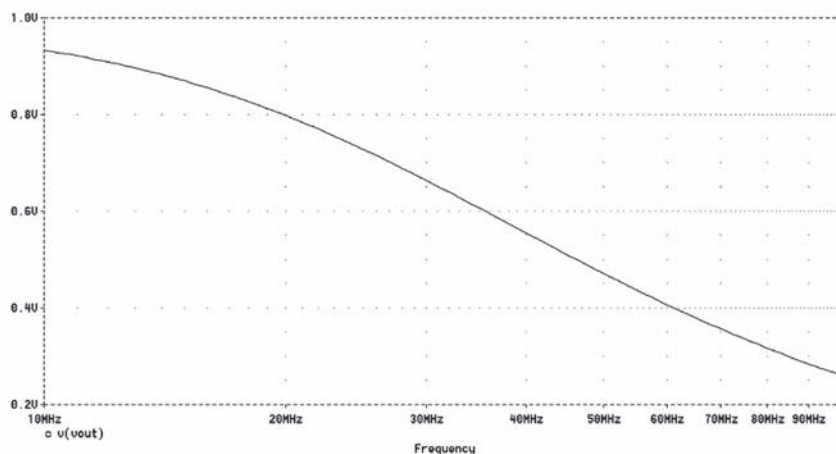


Figure 7-18: Emitter-follower for bootstrapping example, PSPICE result.

We can improve the -3dB bandwidth by bootstrapping the C_{μ} of the input transistor, as shown in **Figure 7-19a**, where we have added a gain of $+1$ across the collector-base junction of Q_1 . The simulation result (**Figure 7-20**) shows that we have improved the bandwidth significantly, at the expense of some gain peaking. One way to reduce the gain peaking is to add some damping (**Figure 7-21**). In **Figure 7-21b**, we show the results (empirically derived) for a damping resistor in the $1\text{ k}\Omega - 10\text{ k}\Omega$ range. It looks like an optimum value for this circuit is $\sim 5\text{ k}\Omega$, which results in good bandwidth with a minimum of gain peaking.

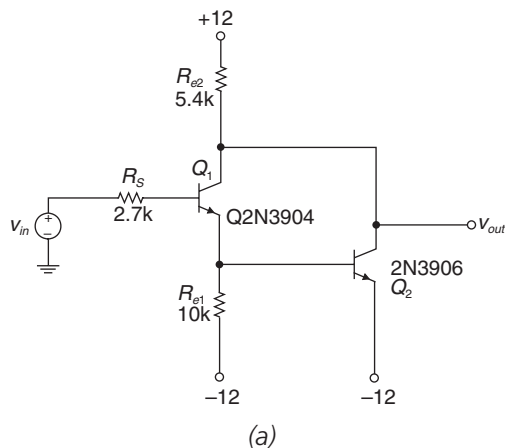
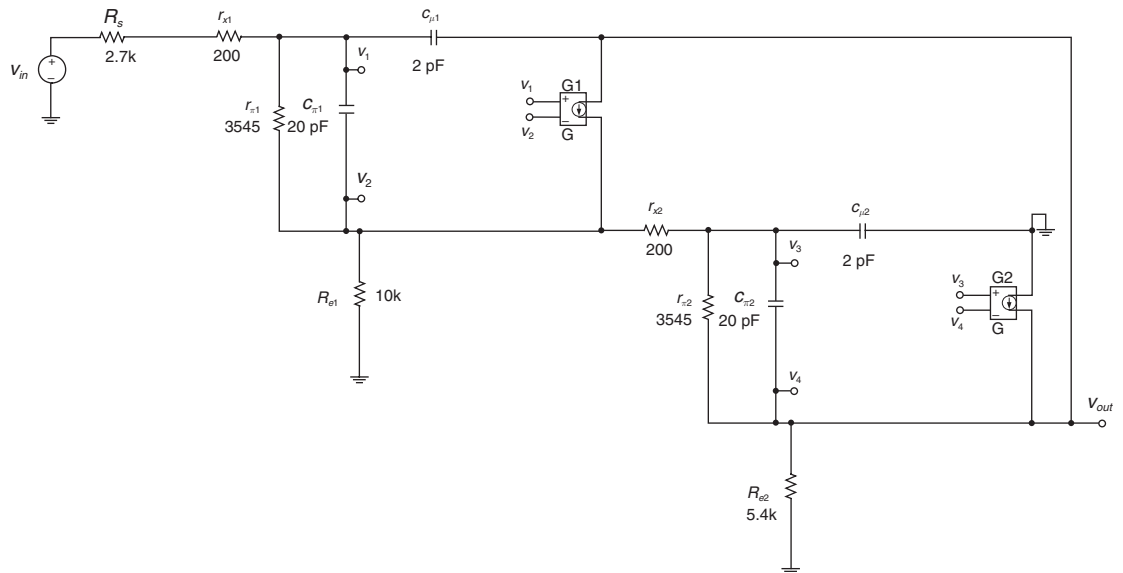


Figure 7-19 (continued on following page):
Bootstrapped emitter-follower. (a) Circuit.



(b)

Figure 7-19 (continued): (b) Small-signal model.

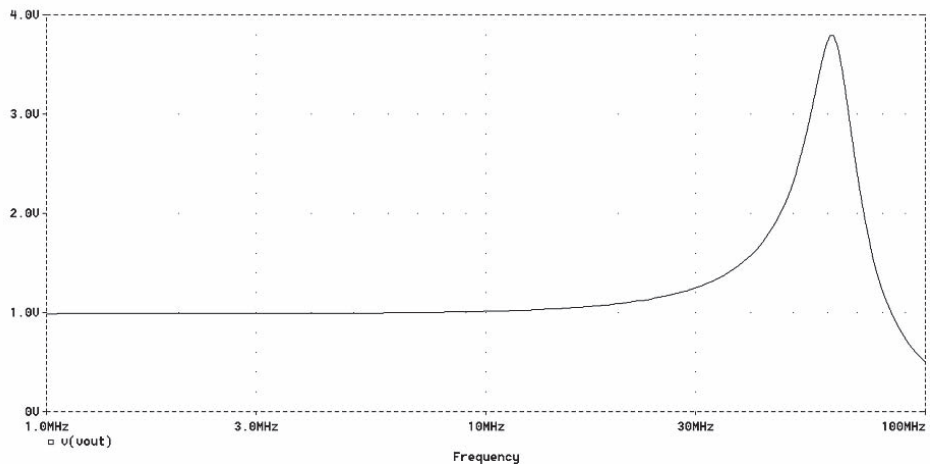
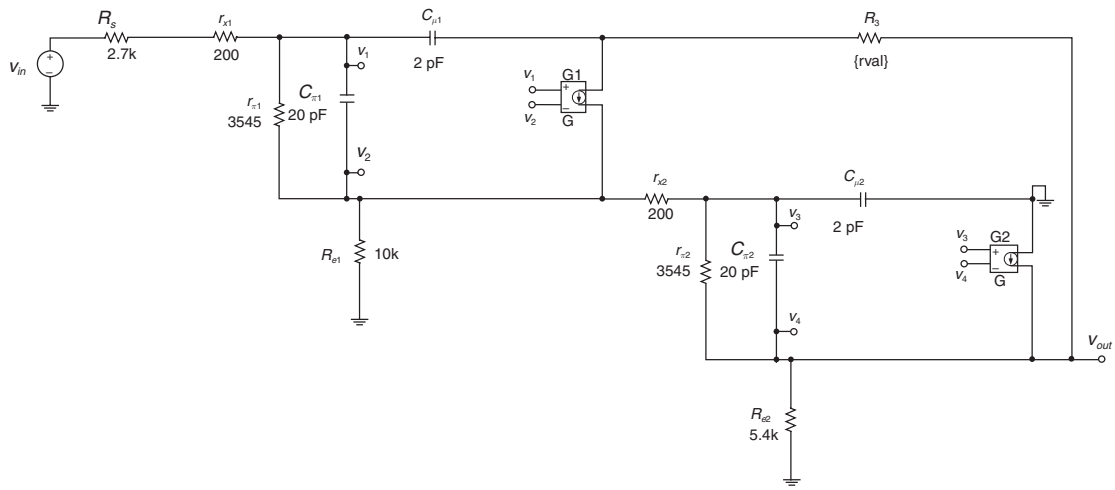
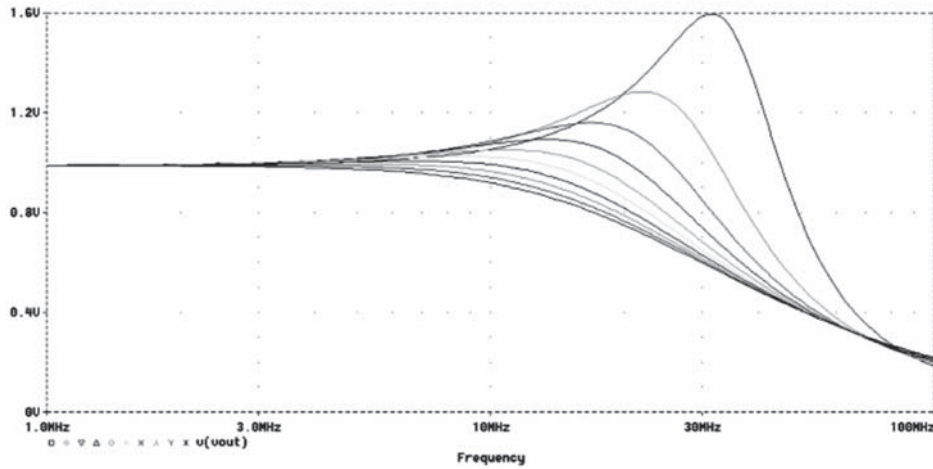


Figure 7-20: Bootstrapped emitter-follower simulation result.



(a)



(b)

Figure 7-21: Bootstrapped emitter-follower with damping added to reduce gain peaking. (a) Circuit small-signal model. (b) PSPICE simulation results for a damping resistor of 1 kΩ, 2 kΩ, ... 10 kΩ.

Example 7.6: Another bootstrapping design example

We'll now revisit an amplifier similar to that from Chapter 6 (**Figure 7-22a**), and further extend the bandwidth by the use of bootstrapping.¹⁵ In this circuit, the dominant time constant is due to the $C_{\mu 3}$ of the input emitter-follower Q_3 interacting with the 2.7-k Ω source resistance. The gain of this circuit is over 100 and the bandwidth (**Figure 7-22b**) is approximately 19 MHz.

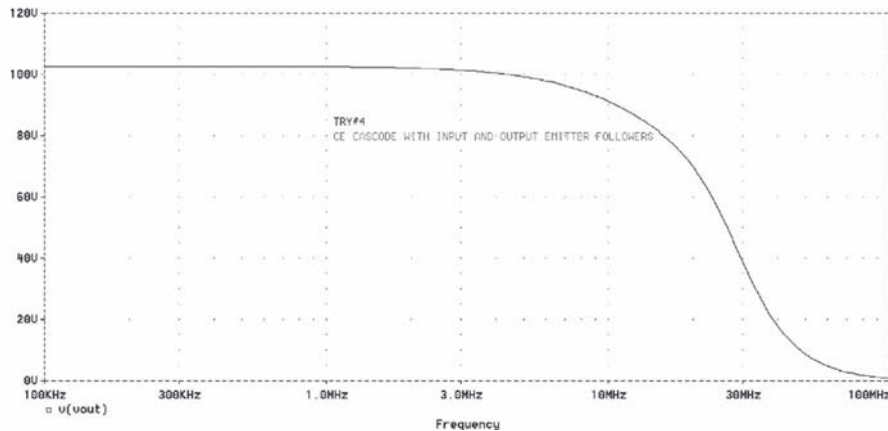
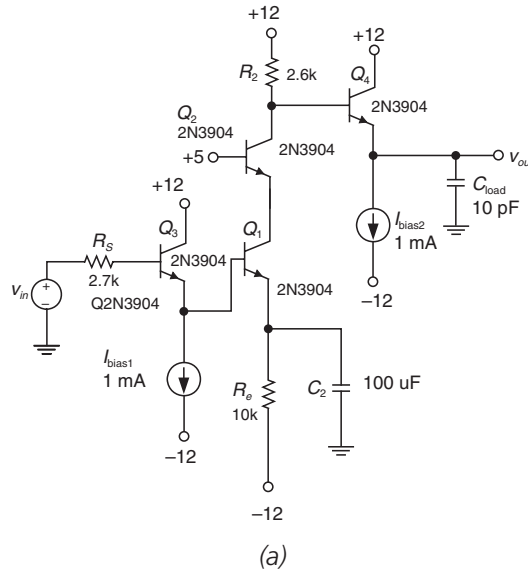


Figure 7-22: Another design example. (a) Circuit.
(b) PSPICE simulation showing a bandwidth of approximately 19 MHz.

¹⁵ This example is similar in topology to that in Chapter 6, but some of the details (bias point levels, etc.) are different. This section resorts exclusively to PSPICE simulations; the open-circuit time constant calculations, although illustrative, are somewhat cumbersome.

In order to reduce the effects of $C_{\mu 3}$, let's force a gain of +1 across the collector-base junction of Q_3 by adding bootstrap transistor Q_5 (**Figure 7-23a**). Q_5 is a PNP emitter-follower, and forces the collector of Q_3 to follow the base of Q_3 incrementally. The PSPICE analysis (**Figure 7-23b**) shows that we have indeed extended the bandwidth of the overall amplifier to approximately 50 MHz, although at the expense of some gain peaking at the high end.

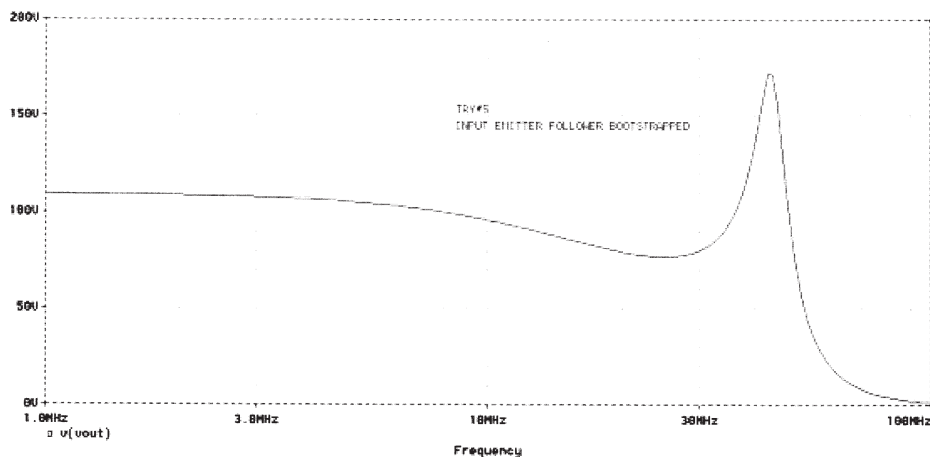
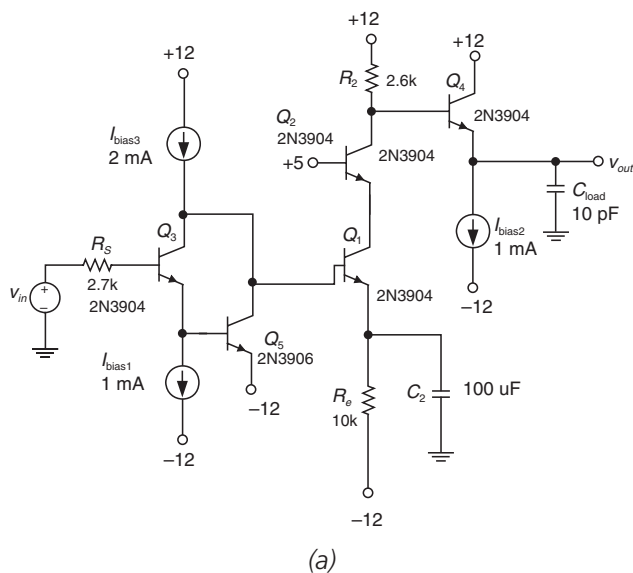


Figure 7-23: Circuit modified by bootstrapping of input emitter-follower Q_3 with bootstrap transistor Q_5 . (a) Circuit. (b) Frequency response from PSPICE.

We can carry this concept one step further by bootstrapping Q_4 (**Figure 7-24a**) by inclusion of bootstrap transistor Q_6 . Note that we get lots of gain peaking (**Figure 7-24b**).

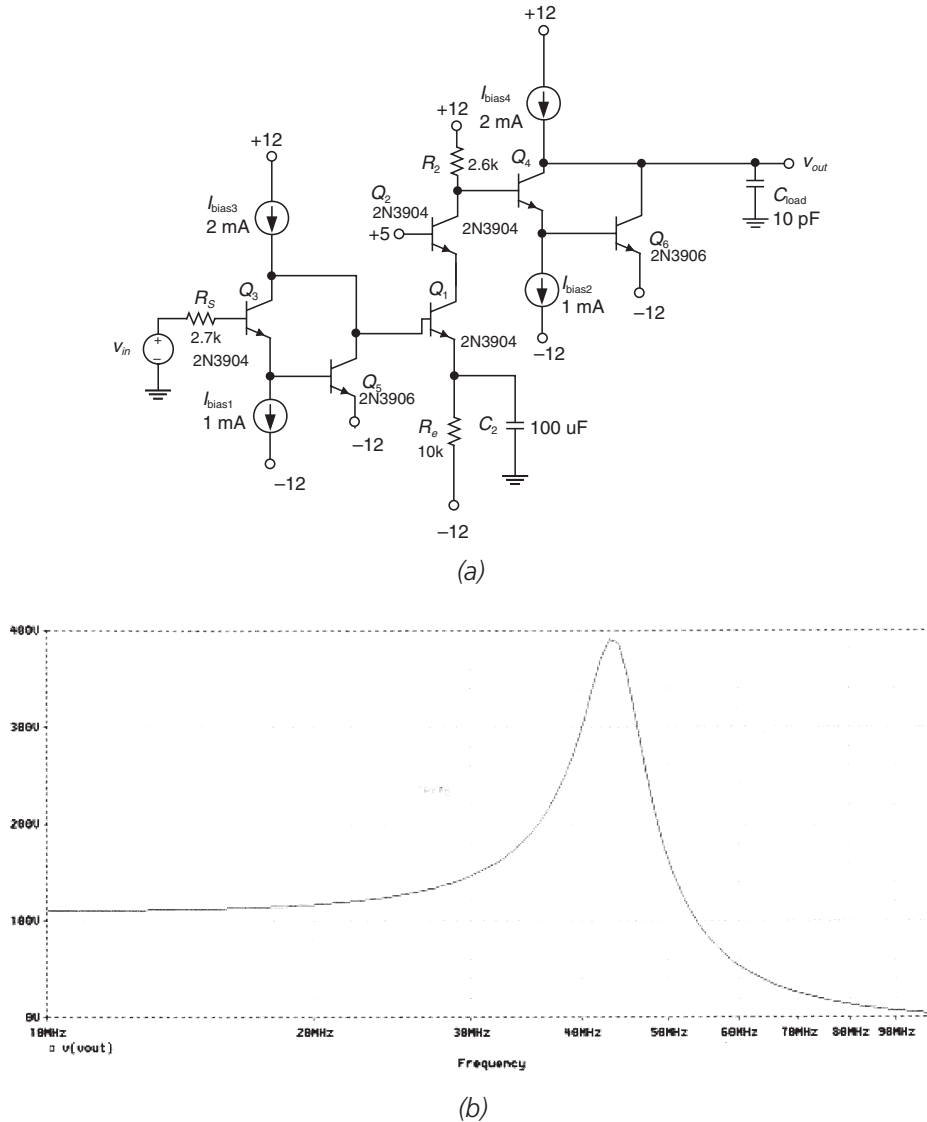


Figure 7-24: Circuit further modified by bootstrapping of output emitter-follower Q_4 . (a) Circuit. (b) Frequency response from PSPICE.

We can qualitatively trace back the gain peaking to several possible causes. We know that gain peaking can occur due to the inductive output impedance of an emitter-follower interacting with any stray capacitance. One strategy to lower the effects of inductive peaking is to change the bias levels of the bootstrap transistors. Shown in **Figure 7-25** is the response of the amplifier with the collector currents of the bootstrap transistors each run at 5 milliamps. We see that the gain peaking has been significantly reduced, while the bandwidth of the overall amplifier is $\sim 80\text{ MHz}$.

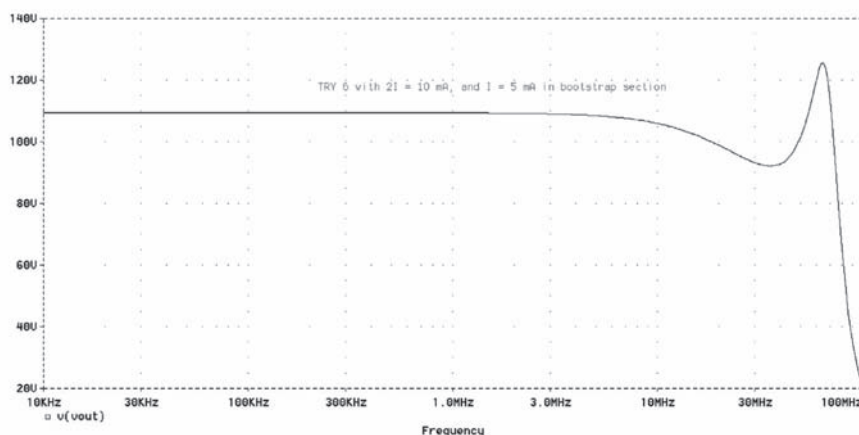


Figure 7-25: Response with emitter-followers run at a collector current of 5 milliamps.

Short-Circuit Time Constants

The method of open-circuit time constants discussed in the previous chapter allows estimation of the high frequency breakpoint f_h of a generic amplifier. A similar method, called *short-circuit time constants*, allows estimation of the low-frequency rolloff of a transistor amplifier due to bypass and coupling capacitors. This same methodology can be applied to a generic circuit to determine the low-frequency breakpoint.

A circuit illustrating the use of short-circuit time constants is shown in **Figure 7-26a**. We know intuitively that the amplifier has the gain-frequency curve of **Figure 7-26b**, for frequencies low enough so that C_π and C_μ don't yet come into play. The method of short-circuit time constants gives us an estimate for the low-frequency -3dB point ω_L .

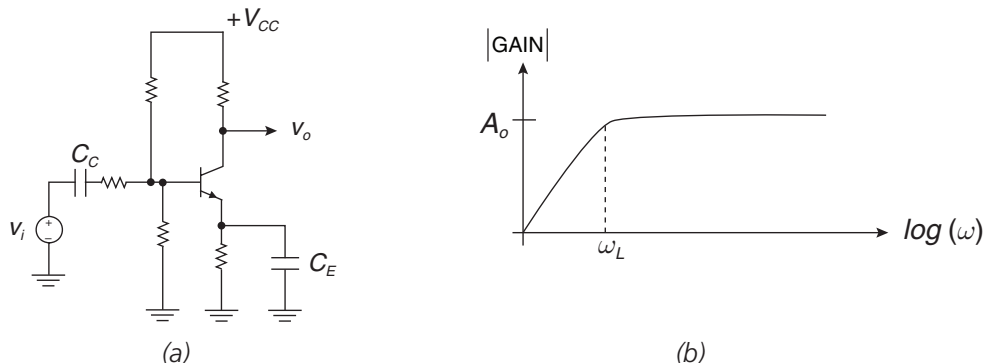


Figure 7-26: Typical AC-coupled amplifier. (a) Circuit showing input coupling capacitor C_C and emitter bypass capacitor C_E . (b) Frequency response.

Let's examine the transfer function of this amplifier, qualitatively at first. If we are concerned with the low-frequency -3dB point, we need not be concerned with C_π and C_μ in the transistors. So, we'll focus on the coupling capacitor C_C and the emitter bypass capacitor C_E .

We know that there's a zero at zero frequency due to the coupling capacitor C_C . This makes sense, since C_C doesn't pass any DC component of the input signal. Secondly, there's a zero at a finite frequency due to the emitter bypass capacitor C_E .¹⁶ A transfer function that captures these functional requirements is:

$$H(s) = K \frac{s(\tau_z s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1)} \quad [7-29]$$

In order to simplify the mathematics,¹⁷ let's assume that the zero is at a very low frequency, so low that we can approximate it as being at zero frequency. This results in the transfer function:

$$H(s) \approx K' \frac{s^2}{(\tau_1 s + 1)(\tau_2 s + 1)} \quad [7-30]$$

If we multiply out the denominator, we find:

$$H(s) \approx K' \frac{s^2}{\tau_1 \tau_2 s^2 + (\tau_1 + \tau_2)s + 1} \quad [7-31]$$

The Bode plot of this approximate transfer function is shown in **Figure 7-27**. We note that the -3dB point is dominated by the higher-frequency pole; in this case, the -3dB point is approximately at frequency $1/\tau_2$. Using a widely spaced pole approximation (which we developed in Chapter 2) we find that the higher frequency pole location can roughly be approximated as:

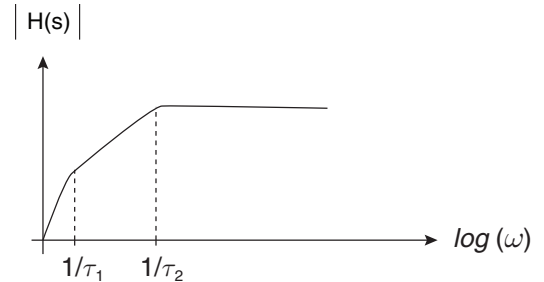
$$p_{\text{high}} \approx -\left(\frac{\tau_1 + \tau_2}{\tau_1 \tau_2}\right) = -\left(\frac{1}{\tau_1} + \frac{1}{\tau_2}\right) \quad [7-32]$$

In the general case with n low-frequency poles (and hence n zeros in the transfer function) we find that the approximation for low-frequency bandwidth ω_L is:

$$\omega_L \approx \sum_{j=1}^n \frac{1}{\tau_{pj}} \quad [7-33]$$

Said in words, an estimate for low frequency bandwidth is equal to the *sum of the reciprocal pole time constants*. Unfortunately, in the general case, it is difficult to calculate these pole time constants.

Figure 7-27: Frequency response of approximate transfer function $H(s)$.



¹⁶ As C_E comes into play, the gain increases because C_E shorts out the emitter resistor R_E . This means C_E contributes a zero to the transfer function at a finite frequency.

¹⁷ This methodology is also done in Gray and Searle's *Electronic Principles*, pp. 542–547. The method is also covered in Gray, et al.'s *Analysis and Design of Analog Integrated Circuits*.

Fortunately, a method for finding the sum of the reciprocal time constants exists. It is the method of short-circuit time constants, which was developed at MIT by Adler and others. Using the method of short-circuit time constants, we find that the sum of the reciprocal time constants is exactly equal to the reciprocal sum of relatively easily calculated short-circuit time constants. Mathematically, we find:

$$\omega_L \approx \sum_{j=1}^N \frac{1}{\tau_{pj}} \approx \sum_{j=1}^N \frac{1}{\tau_{scj}} \quad [7-34]$$

where τ_{sc} are the individual short-circuit time constants. In a method analogous to the method of open-circuit time constants, here is the short-circuit time constant recipe:

- Identify each capacitor that contributes to low-frequency rolloff in your circuit.
- For each capacitor, find the resistance facing this capacitance with other capacitors shorted. (The shorting of the other capacitors is why this is called “short-circuit time constants.”)
- Find the time constant associated with each capacitor.
- Our bandwidth estimate is found by:

$$\omega_L \approx \sum_{j=1}^N \frac{1}{\tau_{scj}} \quad [7-35]$$

Note that we have to be careful in what capacitors we choose in the method of short-circuit-time constants. For instance, in a generic common-emitter amplifier (**Figure 7-28**), we will not use C_π and C_μ in our short-circuit time constant calculation; C_π and C_μ contribute to high-frequency bandwidth limit, but have no contribution to the low-frequency limit.

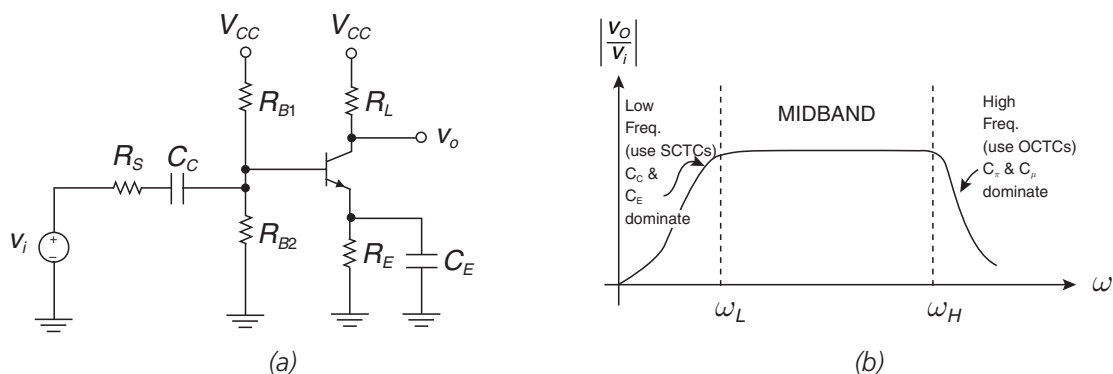


Figure 7-28: Common-emitter amplifier for short-circuit time constants analysis. (a) Circuit. (b) Transfer function, showing low-frequency and high-frequency bandwidth limits. We use open-circuit time constants in the high-frequency model to calculate the upper -3dB point ω_H .

Example 7.7: Short-circuit time constants design example

We'll now do a short-circuit time constants example, and estimate the low-frequency break-point and the midband gain of the circuit of **Figure 7-29a**. Let's assume that the transistor has the following parameters: DC current gain $\beta_F = 100$, base spreading resistance $r_x = 100\Omega$, and small-signal current gain $h_{fe} = 150$. We'll assume that the transistor has $V_{BE} = 0.7$ volts.

The base and collector currents are found by using the circuit of **Figure 7-29b**:

$$\begin{aligned} 6V - I_B (R_{B1} \parallel R_{B2}) - V_{BE} - I_E R_E &= 0 \\ 6V - \left(\frac{I_C}{\beta_F} \right) (R_{B1} \parallel R_{B2}) - V_{BE} - I_C R_E &\approx 0 \\ I_C &\approx \frac{6 - V_{BE}}{\left(R_E + \frac{R_{B1} \parallel R_{B2}}{\beta_F} \right)} \approx \frac{5.3}{\left(1k + \frac{5k}{100} \right)} \approx 5\text{mA} \end{aligned} \quad [7-36]$$

Knowing the bias point value, we can now calculate the small-signal parameters:

$$\begin{aligned} g_m &= \frac{|I_C|}{V_{TH}} = \frac{0.005}{0.026} \approx 0.19 \text{ A/V} \\ r_\pi &= \frac{h_{fe}}{g_m} = \frac{150}{0.19} = 780\Omega \end{aligned} \quad [7-37]$$

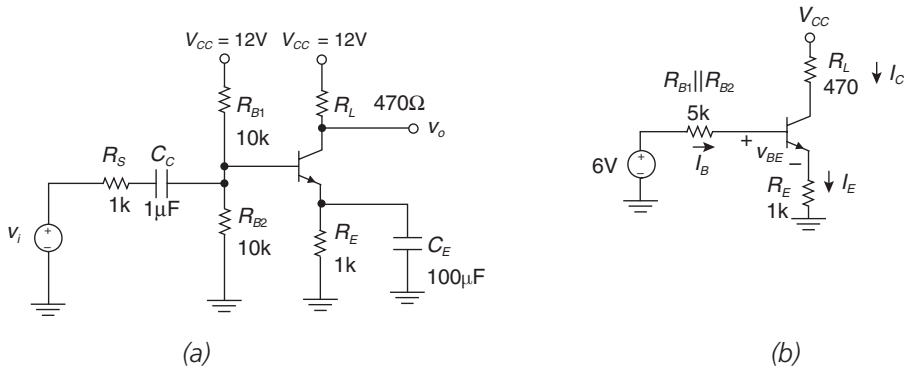


Figure 7-29: Common-emitter amplifier for short-circuit time constants bandwidth estimate. (a) Original circuit. (b) Circuit for finding collector current.

A small-signal model for this circuit, valid for low and midband frequencies, is shown in **Figure 7-30a**. At midband frequencies, the coupling capacitor C_C and the emitter bypass capacitor C_E shorts out, resulting in the circuit of **Figure 7-30b**. We'll now find the midband gain using this circuit:

$$v_x = v_i \left(\frac{(R_{B1} \parallel R_{B2}) \parallel (r_x + r_\pi)}{R_s + (R_{B1} \parallel R_{B2}) \parallel (r_x + r_\pi)} \right) = v_i \left(\frac{5000 \parallel 880}{1000 + 5000 \parallel 880} \right) = v_i \left(\frac{748}{1000 + 748} \right) = 0.428 v_i$$

$$v_\pi = v_x \left(\frac{r_\pi}{r_x + r_\pi} \right) = 0.886 v_x = 0.379 v_i \quad [7-38]$$

$$v_o = -g_m v_\pi R_L = -(0.19)(0.379 v_i)(470) = -33.9 \Rightarrow \frac{v_o}{v_i} = -33.9$$

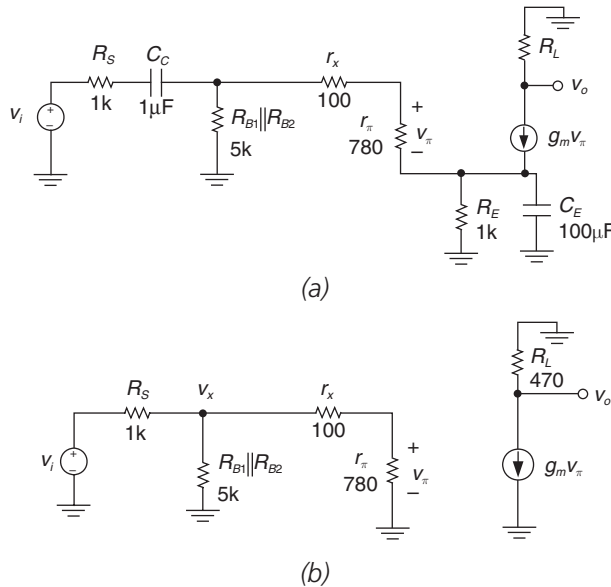


Figure 7-30: Small-signal models of common-emitter amplifier.
 (a) Circuit valid for low and midband frequencies.
 (b) Midband circuit for finding midband gain.

Having found the gain, we'll next estimate the low-frequency breakpoint using short-circuit time constants. The short circuit time constant circuit for the coupling capacitor is shown in **Figure 7-31a**. The short-circuit resistance of the emitter bypass capacitor acting as a short is:

$$R_{SC1} = R_s + (R_{B1} \parallel R_{B2}) \parallel (r_x + r_\pi) = 1000 + 5000 \parallel 880 = 1748\Omega \quad [7-39]$$

The resultant short-circuit time constant for the coupling capacitor is:

$$\tau_{SC1} = R_{SC1} C_C = (1748)(10^{-6}) = 1.75 \times 10^{-3} \text{ sec.} \quad [7-40]$$

We'll next find the short-circuit resistance for the emitter bypass capacitor with the coupling capacitor acting as a short, using the circuit of **Figure 7-31b**. This short-circuit resistance is:

$$R_{SC2} = \frac{R_S \parallel R_{B1} \parallel R_{B2} + r_x + r_\pi}{1 + h_{fe}} = \frac{833 + 100 + 780}{151} = 11.3\Omega \quad [7-41]$$

The resultant short-circuit time constant for the emitter bypass capacitor is:

$$\tau_{SC2} = R_{SC2} C_E = (11.3)(100 \times 10^{-6}) = 1.1 \times 10^{-3} \text{ sec.} \quad [7-42]$$

Our estimate for low-frequency breakpoint is:

$$\omega_L \approx \sum \frac{1}{\tau_{SC}} \approx \frac{1}{\tau_{SC1}} + \frac{1}{\tau_{SC2}} \approx \frac{1}{1.75 \times 10^{-3}} + \frac{1}{1.1 \times 10^{-3}} \approx 1480 \text{ r/sec} \quad [7-43]$$

$$f_L \approx \frac{\omega_L}{2\pi} \approx 235 \text{ Hz}$$

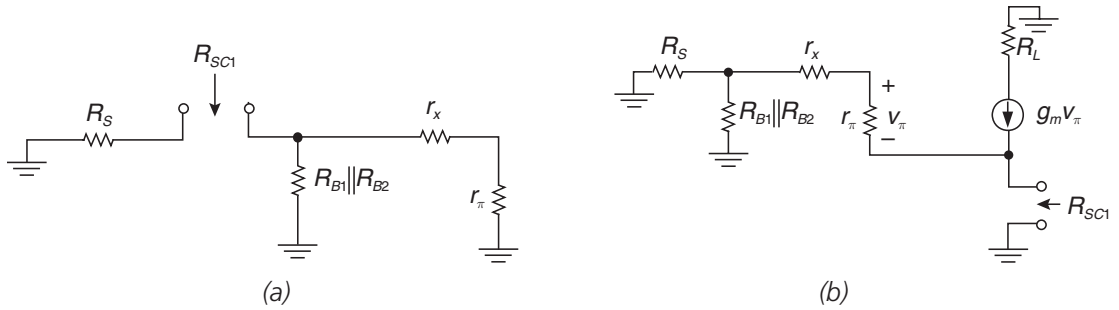


Figure 7-31: Circuits for finding short-circuit time constants. (a) Short-circuit time constant circuit for coupling capacitor C_C . (b) Short-circuit time constant circuit for emitter bypass capacitor C_E .

A PSPICE analysis (**Figure 7-32**) shows that the gain is as calculated and that the -3dB bandwidth is approximately 220 Hz. As in the method of open-circuit time constants, short-circuit time constants give results that are conservative; your actual -3dB breakpoint will be lower than that predicted by the method.

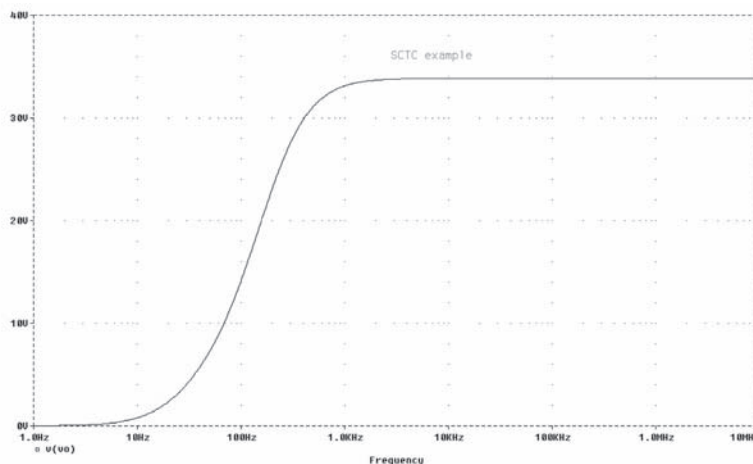


Figure 7-32: PSpice frequency response for short-circuit time constants example showing gain of -34 and -3 dB breakpoint of ~ 220 Hz.

Example 7.8: Peaking amplifier revisited

We'll revisit the peaking amplifier of Chapter 5, but this time without making any restrictions as to the value of transistor base resistance r_x . We saw in Chapter 5 that if the value of r_x is very small, that the transfer function can easily be found in closed form. However, in many small-signal transistors the value of r_x can be significant. For instance, for the 2N3904 transistor the value of r_x is $100\Omega \sim 250\Omega$, depending on collector current bias level. The resultant circuit is shown in **Figure 7-33a**, and the small-signal model is shown in **Figure 7-33b**.

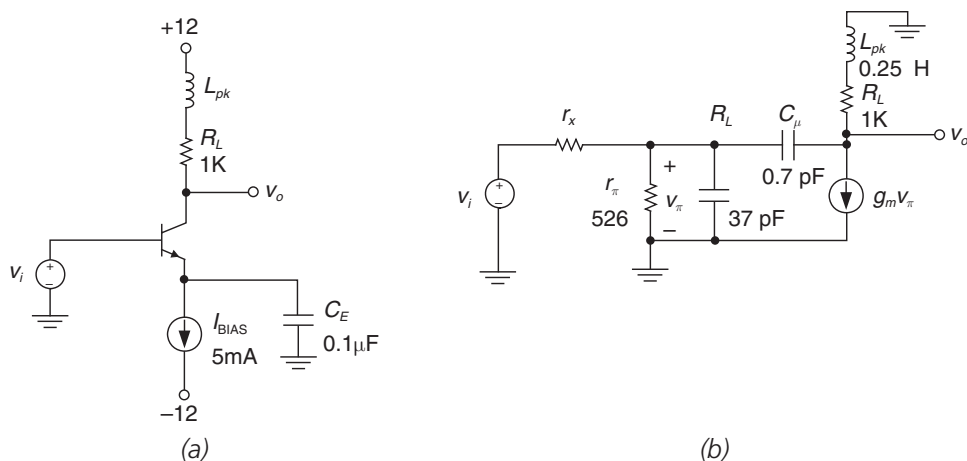


Figure 7-33: Inductively peaked amplifier. (a) Circuit. (b) Small-signal model, including base-spreading resistance r_x .

In Chapter 5, we were able to increase the bandwidth of the peaked amplifier to approximately 390 MHz, by using a peaking inductor $L_{pk} = 0.25$ microhenries. However, in that example we assumed that r_x was zero. Let's examine the performance of the peaking amplifier, but with finite values of base resistance.

In **Figure 7-34**, we see the effect on the inductively peaked amplifier (with $L_{pk} = 0.25$ microhenries) and with $r_x = 0, 10\Omega, 20\Omega, 30\Omega$ and 40Ω . We see that small values of r_x have a significant detrimental effect on bandwidth. Hence, the caveat is to be careful in neglecting r_x in your bandwidth calculations and simulations.¹⁸

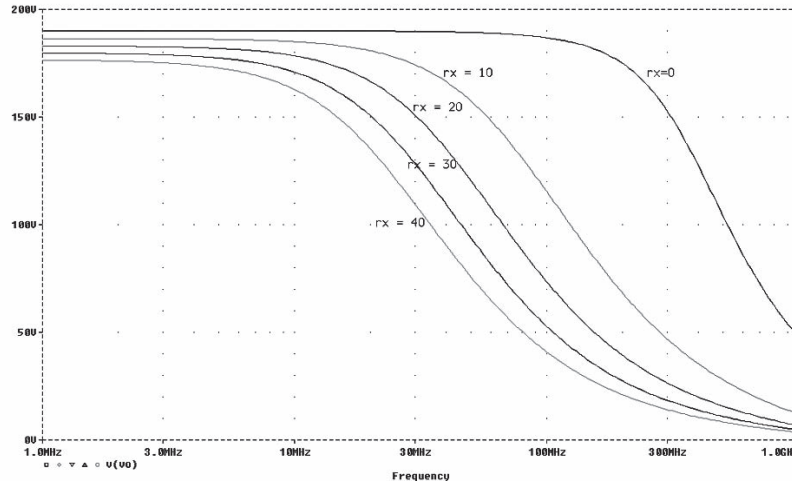


Figure 7-34: PSPICE simulation of inductively peaked ($L_{pk} = 0.25$ microhenries) amplifier, with r_x in the range $0 - 40\Omega$.

Example 7.9: Common-base amplifier

The common-base amplifier is a topology commonly used in high-frequency circuits. In fact, the cascode amplifier used a common-base transistor as an output buffer, as we have seen before. A common-base amplifier is shown in **Figure 7-35a** with $R_E = 1\text{ k}\Omega$ and $R_L = 3.3\text{ k}\Omega$. Let's assume that $V_{cc} = 12\text{V}$ and that $V_{BIAS} = 6\text{V}$. The transistor has $I_C = 1$ milliamp, $h_{fe} = 100$, $f_T = 500\text{ MHz}$, $r_x = 50\Omega$ and $C_\mu = 1\text{ pF}$. The high-frequency small-signal model for this amplifier is shown in **Figure 7-35b**. We'll find the gain and estimate the bandwidth using open-circuit time constants.

¹⁸ Unfortunately, there is very little datasheet information available for most transistors on the value of base spreading resistance. Note that r_x also affects low-frequency gain as well.

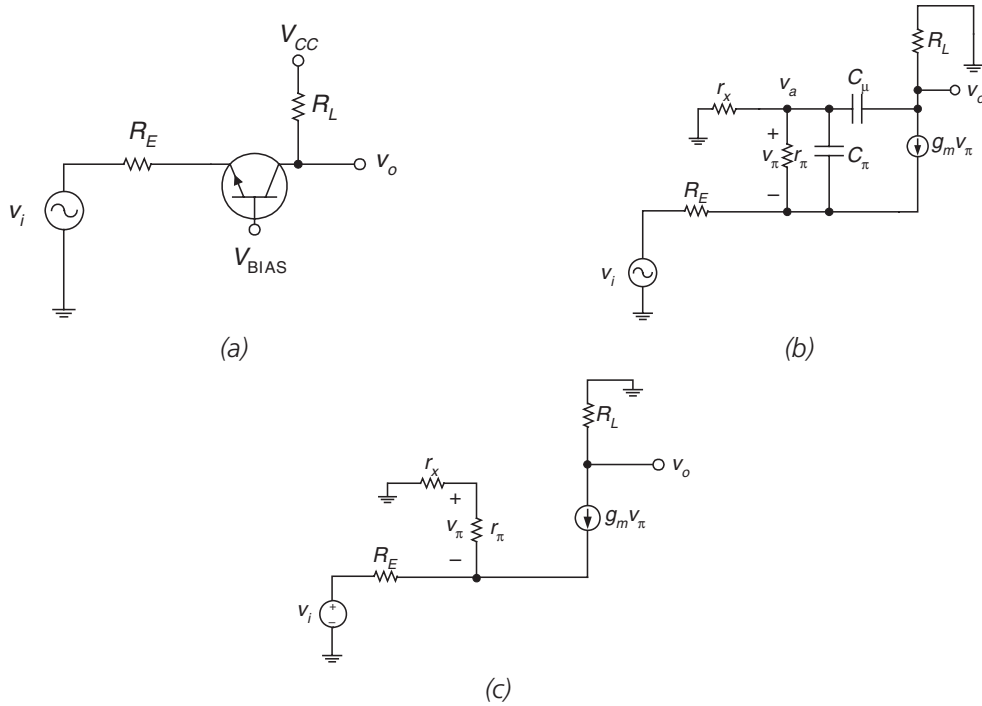


Figure 7-35: Common-base amplifier. (a) Circuit. (b) Small-signal model. (c) Low-frequency small-signal model.

The small-signal parameters for this amplifier are:

$$\begin{aligned}
 g_m &= \frac{|I_c|}{V_{TH}} = \frac{1 \text{ mA}}{26 \text{ mV}} = 0.038 \Omega^{-1} \\
 r_\pi &= \frac{h_{fe}}{g_m} = \frac{100}{0.038 \Omega^{-1}} = 2631 \Omega \\
 C_\pi &= \frac{g_m}{\omega_T} - C_\mu = \frac{0.038}{2\pi(500 \times 10^6)} - 1 \text{ pF} = 11 \text{ pF}
 \end{aligned}
 \tag{7-44}$$

By using the low-frequency small-signal model (**Figure 7-35c**), we can find a closed-form solution for the gain of the common-base amplifier that results in:

$$\frac{v_o}{v_i} \approx \frac{R_L}{R_E + \left(\frac{r_x + r_\pi}{h_{fe}} \right)}
 \tag{7-45}$$

This equation predicts a gain of +3.2 for this amplifier.

We'll predict the bandwidth of the amplifier using open-circuit time constants and the circuits of **Figure 7-36**. For C_{π} , we find the open-circuit time constant:

$$R_{\pi o} = r_{\pi} \left\| \left(\frac{r_x + R_E}{1 + g_m R_E} \right) \right\| \approx \frac{1}{g_m} \approx 26 \Omega \quad [7-46]$$

$$\tau_{\pi o} = R_{\pi o} C_{\pi} = (26)(11 \text{ pF}) = 0.3 \text{ ns}$$

For C_{μ} , we find:

$$\begin{aligned} R_{EQ1} &= r_x \left\| \left(r_{\pi} + (1 + h_{fe}) R_E \right) \right\| \approx r_x \approx 50 \Omega \\ G_M &\approx \frac{g_m}{1 + g_m R_E} \approx \frac{1}{R_E} \approx 0.001 \\ R_{\mu o} &= R_{EQ1} + R_L + G_M R_L R_{EQ1} = 50 + 3300 + (0.001)(3300)(50) \approx 3515 \Omega \\ \tau_{\mu o} &= R_{\mu o} C_{\mu} = (3515)(1 \text{ pF}) = 3.5 \text{ ns} \end{aligned} \quad [7-47]$$

The sum of open-circuit time constants for this amplifier is 3.8 nanoseconds, and hence our estimate of bandwidth is:

$$\omega_h \approx \frac{1}{3.8 \text{ ns}} \approx 263 \text{ Mrad/sec} \quad [7-48]$$

$$f_h = \frac{\omega_h}{2\pi} \approx 41.9 \text{ MHz}$$

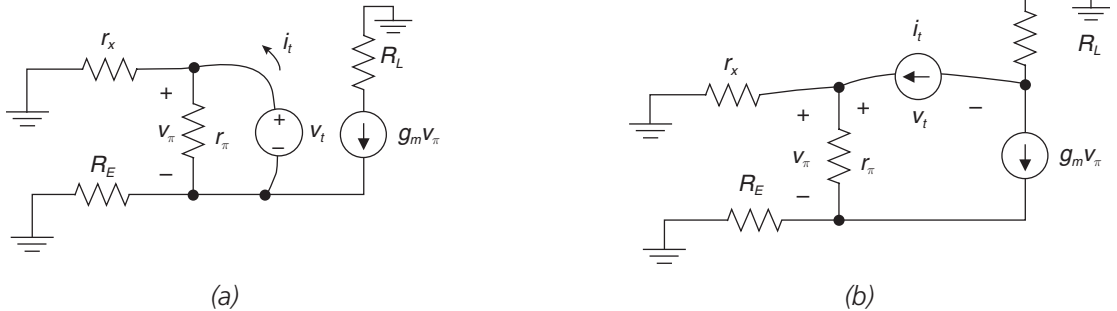


Figure 7-36: Circuits for finding OCTCs for common-base amplifier.
(a) Circuit for C_{π} . (b) Circuit for C_{μ} .

PSPICE (**Figure 7-37**) shows that the gain is +3.2 and the -3dB bandwidth is 44.4 MHz.

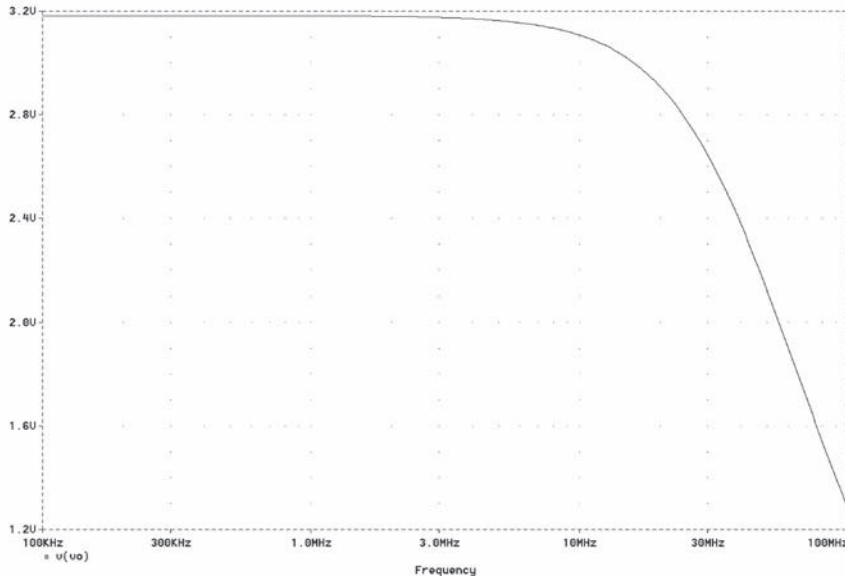


Figure 7-37: PSPICE result for common-base amplifier, showing a gain of +3.2 and a -3dB bandwidth of 44.4 MHz.

Example 7.10: Current amplifier

An amplifier used for fast current switching is shown in **Figure 7-38**. Circuits similar to this, where currents are switched and/or amplified, are used extensively in high-frequency circuits such as current-feedback operational amplifiers. We'll estimate the bandwidth of this circuit assuming small signal¹⁹ operation, using open-circuit time constants. For incremental differential-mode signals, we can ground the emitter connection (as the incremental voltage swing at this node is zero), resulting in the circuits shown. By inspection, the open circuit resistances for C_π and C_μ are the same, resulting in:

$$\begin{aligned} R_{o1} &= r_x \parallel r_\pi \\ R_{o2} &= r_x \parallel r_\pi \\ \sum \tau_{oc} &= (r_x \parallel r_\pi)(C_\pi + C_\mu) \\ \omega_h &\approx \frac{1}{\sum \tau_{oc}} = \frac{1}{(r_x \parallel r_\pi) + (C_\pi + C_\mu)} \end{aligned} \quad [7-49]$$

This is likely to be at a very high frequency since there is no Miller effect. For example, for a 2N3904 switch with $f_T = 300 \text{ MHz}$, $\omega_T = 1.89 \times 10^9 \text{ radians/second}$, $r_x = 100\Omega$, $h_{fe} = 150$, $C_\mu = 2 \text{ pF}$. If we bias the transistors with $I_{BIAS} = 4 \text{ milliamps}$, then $I_{c1} = I_{c2} = 2 \text{ milliamps}$. We

¹⁹ We'll see more on current-feedback opamps in a later chapter.

then find that $C_\pi = 38.8$ pF, and $r_\pi = 1950\Omega$. This results in an open-circuit time constant estimate of bandwidth $\omega_h \approx 257$ Mrad/sec ($f_h \approx 41$ MHz). SPICE (**Figure 7-39**) shows $f_h = 41.5$ MHz, so the estimate is pretty good.

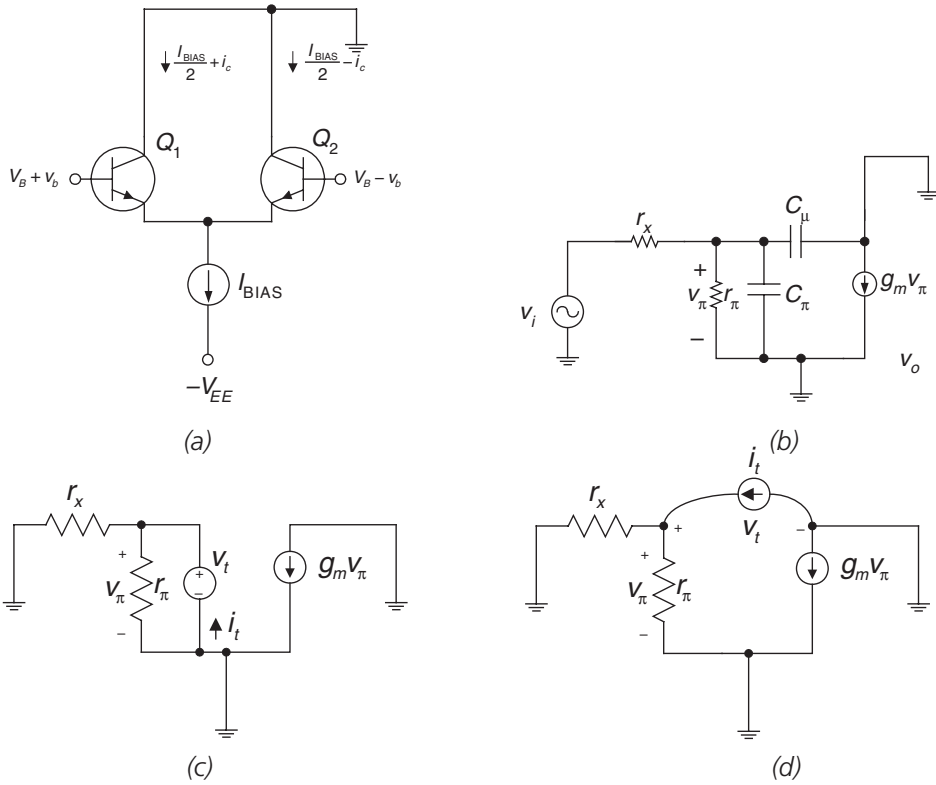


Figure 7-38: Current switching amplifier. (a) Circuit. (b) High-frequency model. (c) Circuit for finding C_π OCTC. (d) Circuit for finding C_μ OCTC.

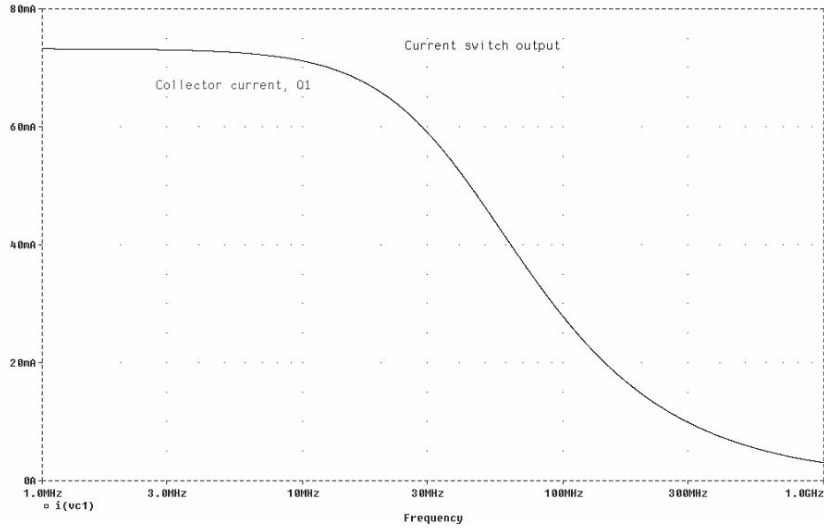


Figure 7-39: Current switching amplifier frequency response, showing a -3dB bandwidth of 41.5 MHz.

Pole Splitting

Pole splitting is a technique widely used to tailor the response of amplifiers, especially in operational amplifiers. We'll illustrate pole splitting, using the circuit of **Figure 7-40**. This circuit models an amplifier with gain $-A$, input resistance R_i , input capacitance C_i , output resistance R_o and output capacitance C_o . We add a feedback capacitor C_f to tailor the pole locations, with the following explanation. It may be counter-intuitive, but there are two poles in this system (not three poles).

First, using the circuit, we'll write the node equations for nodes v_x and the output node v_o :

$$\begin{aligned} (1) & (v_i - v_x)G_i - v_x C_i s + (v_o - v_x)C_f s = 0 \\ (2) & (-A v_x - v_o)G_o - v_o C_o s + (v_x - v_o)C_f s = 0 \end{aligned} \quad [7-50]$$

Let's next group terms in v_x and v_o :

$$\begin{aligned} (1) & -v_x [G_i + (C_i + C_f)s] + v_o C_f s = -v_i G_i \\ (2) & v_x [C_f s - A G_o] - v_o [G_o + (C_o + C_f)s] = 0 \end{aligned} \quad [7-51]$$

Next, divide (1) by G_i and divide (2) by G_o :

$$\begin{aligned} (1) & -v_x [1 + R_i (C_i + C_f)s] + v_o R_i C_f s = -v_i \\ (2) & v_x [R_o C_f s - A] - v_o [1 + R_o (C_o + C_f)s] = 0 \end{aligned} \quad [7-52]$$

We put this into matrix form in anticipation of using Cramer's rule as follows:

$$\begin{bmatrix} R_i(C_i + C_f)s + 1 & R_i C_f s \\ R_o C_f s - A & R_o(C_o + C_f)s + 1 \end{bmatrix} \begin{bmatrix} v_x \\ v_o \end{bmatrix} = \begin{bmatrix} -v_i \\ 0 \end{bmatrix} \quad [7-53]$$

Using Cramer's rule, we find the solution for the output voltage v_o :

$$v_o = \frac{\det \begin{bmatrix} R_i(C_i + C_f)s + 1 & -v_i \\ R_o C_f s - A & 0 \end{bmatrix}}{\det \begin{bmatrix} R_i(C_i + C_f)s + 1 & R_i C_f s \\ R_o C_f s - A & R_o(C_o + C_f)s + 1 \end{bmatrix}} \quad [7-54]$$

After much algebraic manipulation, we find the solution for the input-output transfer function:

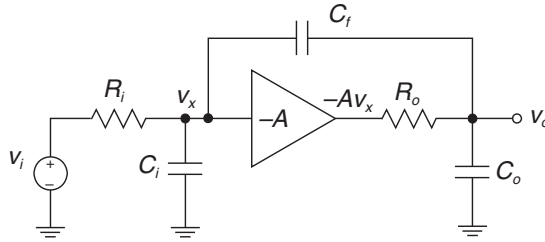
$$\frac{v_o}{v_i} = \frac{-A \left(1 - \frac{R_o C_f s}{A} \right)}{R_i R_o (C_i C_o + C_i C_f + C_f C_o) s^2 + (R_i (C_i + (1 + A) C_f) + R_o (C_o + C_f)) s + 1} \quad [7-55]$$

Let's next do a sanity check with $C_f = 0$ and determine if the transfer function looks correct:

$$\left. \frac{v_o}{v_i} \right|_{C_f=0} = \frac{-A}{R_i R_o C_i C_o s^2 + (R_i C_i + R_o C_o) s + 1} = \frac{-A}{(R_i C_i s + 1)(R_o C_o s + 1)} \quad [7-56]$$

With $C_f = 0$, the input and output circuits are uncoupled. This is correct, since with $C_f = 0$ we have one pole at $1/(R_i C_i)$ and a second pole at $1/(R_o C_o)$.

Figure 7-40: Circuit used to illustrate pole splitting.



In another sanity check, we can apply open-circuit time constants to this amplifier, and determine whether the open-circuit time constant results sum to the same value as the coefficient of the s term in the gain expression earlier derived.

For C_i , we open-circuit C_o and C_f , and find the resistance facing C_i using the circuit of **Figure 7-41a**.

$$\begin{aligned} R_{o1} &= R_i \\ \tau_{o1} &= R_{o1} C_i = R_i C_i \end{aligned} \quad [7-57]$$

Chapter 7

For C_o , we open-circuit C_i and C_f and find the resistance facing C_o using the circuit of **Figure 7-41b**.

$$\begin{aligned} R_{o2} &= R_o \\ \tau_{o2} &= R_{o2}C_o = R_oC_o \end{aligned} \quad [7-58]$$

For C_f , we need to do a little bit more work using the circuit of **Figure 7-41c**.

$$\begin{aligned} v_t &= i_t R_i + i_t R_o + A i_t R_i \\ R_{o3} &= \frac{v_t}{i_t} = R_o + (1 + A) R_i \\ \tau_{o3} &= R_{o3} C_f = (R_o + (1 + A) R_i) C_f \end{aligned} \quad [7-59]$$

The sum of the open-circuit time constants is:

$$\sum \tau_{oc} = \tau_{o1} + \tau_{o2} + \tau_{o3} = R_i C_i + R_o C_o + (R_o + (1 + A) R_i) C_f \quad [7-60]$$

This is exactly the same as the coefficient of the s term in the closed-form solution, and this is a further sanity check on the result of Eq. 7-55.

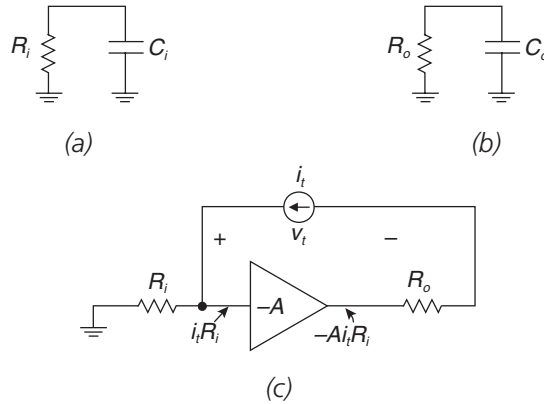


Figure 7-41: Open-circuit time constants circuits for pole split amplifier.

(a) OCTC circuit for C_i . (b) OCTC circuit for C_o . (c) OCTC circuit for C_f .

In order to find the pole locations, we'll assume that we have two poles and that they are widely spaced and on the negative real axis, a circumstance that occurs often in wideband amplifiers. Our widely spaced pole location approximation for a second-order system is:

$$\begin{aligned} H(s) &= \frac{1}{a_2 s^2 + a_1 s + 1} \\ p_{low} &\approx -\frac{1}{a_1} \\ p_{high} &\approx -\frac{a_1}{a_2} \end{aligned} \quad [7-61]$$

Using this approximation, we find the approximate low-frequency (p_{low}) and high-frequency pole (p_{high}) locations for this amplifier as:

$$p_{low} \approx -\frac{1}{(R_i(C_i + (1+A)C_f) + R_o(C_o + C_f))} \quad [7-62]$$

$$p_{high} \approx -\frac{(R_i(C_i + (1+A)C_f) + R_o(C_o + C_f))}{R_i R_o (C_i C_o + C_i C_f + C_f C_o)}$$

We can further approximate if we assume that $AR_i C_f \gg R_i C_i$, $R_o C_o$ and $R_o C_f$:

$$p_{low} \approx -\frac{1}{AR_i C_f}$$

$$p_{high} \approx -\frac{AR_i C_f}{R_i R_o (C_i C_o + C_i C_f + C_f C_o)} \approx -\frac{A}{R_o \left(\frac{C_i C_o}{C_f} + C_i + C_o \right)} \quad [7-63]$$

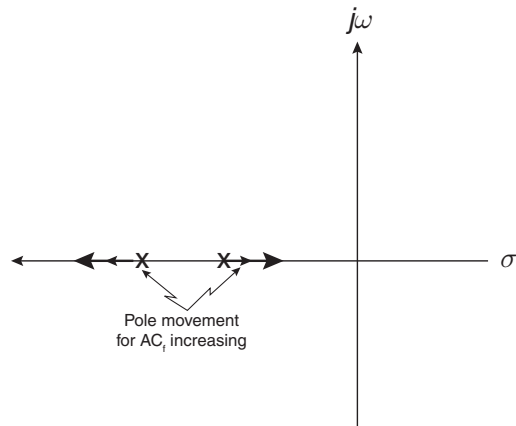
We note the important functional relationship: as AC_f increases, the low-frequency pole decreases in frequency, and the high-frequency pole increases in frequency, as shown in **Figure 7-42**. In fact, the dominant pole is due to the input resistance and Miller multiplication of the feedback capacitance. In op-amps, it's useful to have a single dominant pole, and pole-splitting is used to accomplish this.

Another model of a pole-split amplifier is shown in **Figure 7-43**. This models a common-emitter amplifier driving a capacitive load. We can start out finding the transfer function by writing the node equations at the v_a and v_o nodes:

$$(1) (v_i - v_a)G_i - v_a C_i s + (v_o - v_a)C_f s = 0$$

$$(2) -g_m v_a - v_o G_o - v_o C_o s + (v_a - v_o)C_f s = 0 \quad [7-64]$$

Figure 7-42: Movement of poles as AC_f increases.



Chapter 7

We note that these node equations are identical to that found in the previous pole-split amplifier. Therefore, we can use the previous result with the following substitution:

$$A \Rightarrow g_m R_o \quad [7-65]$$

Therefore, the transfer function for this amplifier is:

$$\frac{v_o}{v_i} = \frac{-g_m R_o \left(1 - \frac{C_f s}{g_m}\right)}{R_i R_o (C_i C_o + C_i C_f + C_f C_o) s^2 + \left(R_i (C_i + (1 + g_m R_o) C_f) + R_o (C_o + C_f)\right) s + 1} \quad [7-66]$$

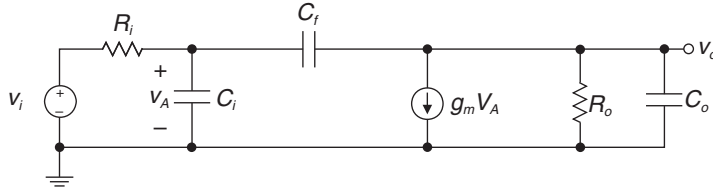


Figure 7-43: Another pole-split amplifier. This model is typical of that found in single-stage transistor amplifiers. The equations show that this amplifier is topologically equivalent to the previously shown pole split amplifier.

Example 7.11: Pole splitting

A numerical example will show how pole splitting works. Let's consider the amplifier of **Figure 7-43** with the following parameters: $g_m = 0.01$ A/V; $R_o = 10^5 \Omega$; $R_i = 10^6 \Omega$; $C_i = C_o = 10$ pF. We'll see how the pole locations move as C_f is varied from zero to 30 picofarads. With $C_f = 0$, we find two uncoupled poles as follows:

$$p_1 = -\frac{1}{R_i C_i} = -10^5 \text{ rad/sec}$$

$$p_2 = -\frac{1}{R_o C_o} = -10^6 \text{ rad/sec} \quad [7-67]$$

When we increase C_f , the low-frequency pole moves down in frequency, and the high-frequency pole moves up in frequency. We can approximate the pole locations if we assume that $g_m R_o R_i C_f \gg R_i C_i$, $R_o C_o$ and $R_o C_f$ as follows:

$$p_{low} \approx -\frac{1}{g_m R_o R_i C_f}$$

$$p_{high} \approx -\frac{g_m}{\left(\frac{C_i C_o}{C_f} + C_i + C_o\right)} \quad [7-68]$$

The plot of **Figure 7-44** shows the pole locations as C_f increases.

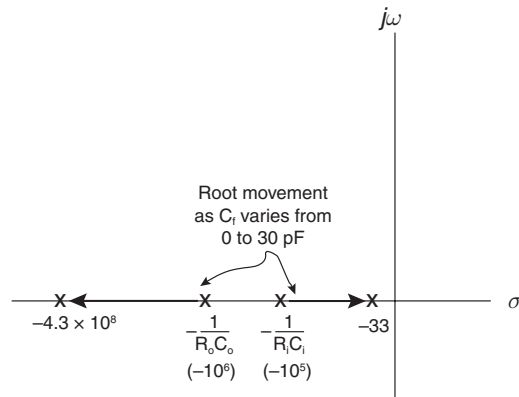


Figure 7-44: Movement of poles in pole-split amplifier example as C_f is varied from zero to 30 picofarads.

Table 7-2: Pole and zero movement in pole-split amplifier as a function of feedback capacitor C_f .

C_f	ω_{p1}	ω_{p2}	ω_z
0	-105 rad/sec.	-106 rad/sec.	$+\infty$
1 pF	-988	-8.4×10^7	+1010
5 pF	-199	-2.5×10^8	$+2 \times 10^9$
10 pF	-100	-3.3×10^8	+109
20 pF	-49	-4×10^8	$+5 \times 10^8$
30 pF	-33	-4.3×10^8	$+3.3 \times 10^8$

Chapter 7 Problems

Problem 7.1

- Assume that the circuit in **Figure 7-45** is driven by an AC voltage source. Using an intuitive approach, thought experiments, etc. sketch the shape of the Bode magnitude response of v_o/v_i .
- Using the method of short-circuit time constants, find the low-frequency breakpoint f_L .

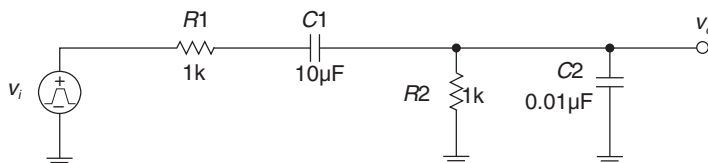
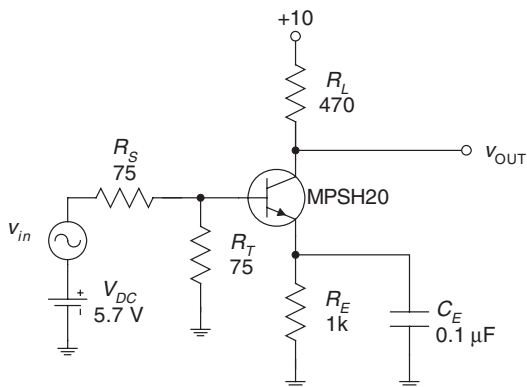


Figure 7-45: Circuit for Problem 7.1.

Problem 7.2

For the transistor amplifier shown in **Figure 7-46**, estimate the low-frequency -3dB point using the method of short-circuit time constants. For the MPSH20 transistor you may assume $r_x = 20\Omega$, AC beta $h_{fe} \approx$ DC beta $h_{FE} = 25$, $C_\mu = 0.9\text{ pF}$ and $f_T = 630\text{ MHz}$. Make reasonable approximations, state them, and justify them. Assume room temperature is 25°C and hence $kT/q = 26\text{ millivolts}$.

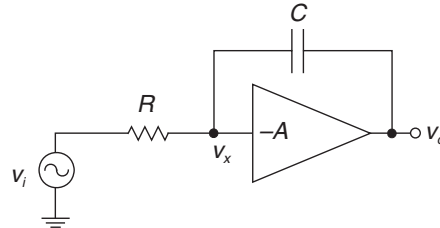
Figure 7-46: Common-emitter amplifier for Problem 7.2.



Problem 7.3

Shown in **Figure 7-47** is a possible model for an amplifier exhibiting the Miller effect. The amplifier has infinite input impedance, zero output impedance, gain $-A$ with transfer function $v_o = -Av_x$. Using the method of open-circuit time constants, find the -3dB bandwidth of this amplifier. What is the bandwidth if $A = +1$?

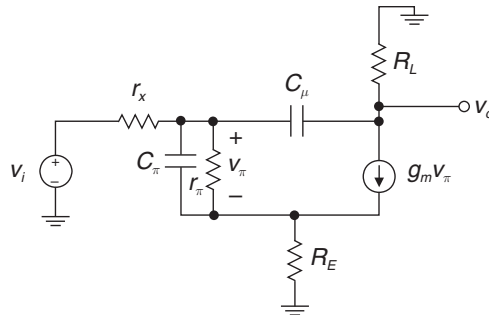
Figure 7-47: Circuit for Problem 7.3.



Problem 7.4

Shown in **Figure 7-48** is the small-signal model of a common-emitter amplifier with emitter degeneration. The circuit has load resistor $R_L = 1\text{ k}\Omega$ and emitter degeneration resistor $R_E = 22\Omega$. Assume that the transistor is biased at $I_C = 10$ milliamps. Find the gain and estimate the bandwidth using open-circuit time constants. Small-signal parameters are: $r_x = 100\Omega$; $r_\pi = 1\text{ k}\Omega$; $g_m = 0.1\text{ A/V}$; $C_\pi = 20\text{ pF}$; $C_\mu = 2\text{ pF}$.

Figure 7-48: Common-emitter amplifier with emitter degeneration for Problem 7.4 and Problem 7.5.



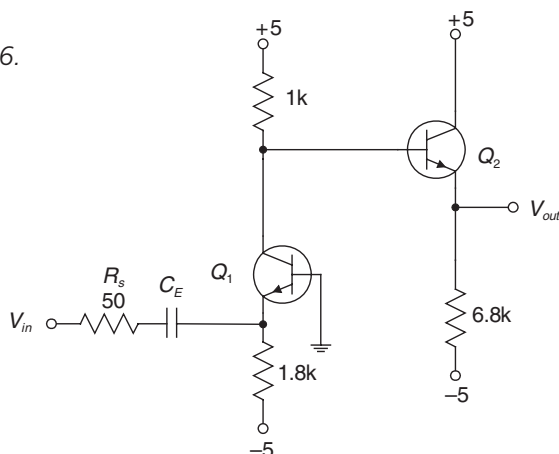
Problem 7.5

Simulate the circuit in **Figure 7-48** using PSPICE. In this simulation, comment on the differences between circuit operation for $R_E = 22\Omega$, 47Ω and 100Ω .

Problem 7.6

The following circuit is a video amplifier that uses a common-base voltage amplifier driving an emitter-follower. The source resistance R_s may be, for instance, the resistance of a 50Ω transmission line used to input the signal to the amplifier.

Figure 7-49: Circuit for Problem 7.6.



You may assume that, under the operating conditions in the circuit, $h_{FE} = 50$, $h_{fe} = 50$ and $r_x = 50\Omega$ for both transistors.

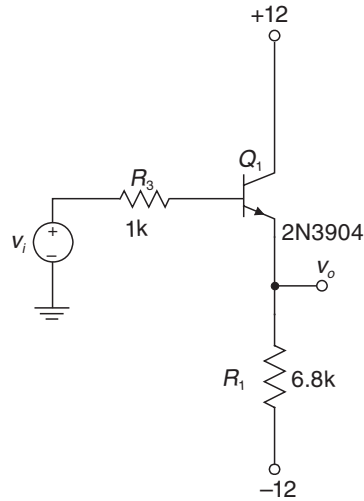
- What is the operating point value of the output voltage V_{out} ?
- What is the midband voltage gain v_o/v_{in} ?
- Assume that the design specifies a lower cutoff frequency of 50 Hz. What value of C_E is required?
- Verify your results using SPICE.

Problem 7.7

For the circuit in **Figure 7-50**, assume that the input voltage source is an AC source with a DC value of 0 volts and that the circuit operates at 25°C.

- Find the low frequency and midband gain of the emitter-follower, assuming that transistor internal capacitances C_π and C_μ have no effect.
- Using the method of open-circuit time constants, estimate the bandwidth of the buffer shown below. (In your OCTC estimate, find appropriate values of C_π and C_μ from the 2N3904 datasheet.)
- Simulate your circuit using SPICE and compare your OCTC result with the SPICE result. In your SPICE model, directly input the small-signal model (i.e., don't use the SPICE 2N3904 model; input directly C_π , C_μ , r_π , the $g_m v_\pi$ generator, etc. in a small-signal model).

Figure 7-50: Circuit for Problem 7.7.



References

- Abidi, A., "On the operation of cascode gain stages," *IEEE Journal of Solid-State Circuits*, vol. 23, no. 6, December 1988, pp. 1434–1437. *An interesting paper illustrating the use of bootstrapping to improve the DC gain of a MOS amplifier.*
- Barna, A., "On the Transient Response of Emitter Followers," *IEEE Journal of Solid-State Circuits*, June 1973, pp. 233–235.
- Centurelli, F., Luzzi, R., Olivieri, M., and Trifiletti, A., "A bootstrap technique for wideband amplifiers," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 48, no. 10, October 2002, pp. 1474–1480.
- Chuang, C. T., "Analysis of the settling behavior of an operational amplifier," *IEEE Journal of Solid-State Circuits*, vol. 17, no. 1, February 1982, pp. 74–80.
- Choma, J., Jr., "Simplified design guidelines for dominant pole amplifiers peaked actively by emitter or source followers," *IEEE Transactions on Circuits and Systems*, vol. 36, no. 7, July 1989, pp. 1005–1010.
- Eschauzier, R. G. H., Kerklaan, L. P. T., and Huijsing, J. H., "A 100 MHz 100 dB operational amplifier with multipath nested miller compensation structure," *Proceedings of the 1992 IEEE Solid-State Circuits Conference*, (ISSCC), February 19–21, 1992, pp. 196–197.

- Filipkowski, A., "Poles and zeros in transistor amplifiers introduced by Miller effect," *IEEE Transactions on Education*, vol. 42, no. 4, November 1999, pp. 349–351.
- Fong, Keng Leong, and Meyer, R. G., "High-frequency nonlinearity analysis of common-emitter and differential-pair transconductance stages," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 4, April 1998, pp. 548–555.
- Grebene, Alan, *Bipolar and MOS Analog Integrated Circuit Design*, John Wiley, 1984.
- Gray, Paul R., Hurst, Paul, Lewis, Stephen, and Meyer, Robert, *Analysis and Design of Analog Integrated Circuits*, 4th edition, John Wiley, 2001.
- Hamilton, D. K., "Use of inductive compensation for improving bandwidth and noise performance of high frequency optical receiver preamplifiers," *IEEE Proceedings G: Circuits, Devices and Systems*, vol. 138, no. 1, Feb. 1991, pp. 52–55.
- Ki, Wing-Hung, Der, L., and Lam, S., "Re-examination of pole splitting of a generic single stage amplifier," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 44, no. 1, January 1997, pp. 70–74.
- Kozikowski, J., "Analysis and Design of Emitter Followers at High Frequencies," *IEEE Transactions on Circuit Theory*, 1964, pp. 129–136.
- Lee, Thomas, *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge University Press, 1998.
- Makris, C. A., and Toumazou, C., "Two pole, high speed operational amplifier modelling, methods and techniques," *Proceedings of the European Conference on Circuit Theory and Design*, September 5–8, 1989, pp. 304–308.
- Oh, Yong-Hun, and Lee, Sang-Gug, "An inductance enhancement technique and its application to a shunt-peaked 2.5 Gb/s transimpedance amplifier design," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 51, no. 11, Nov. 2004, pp. 624–628.
- Palmisano, G., and Palumbo, G., "An optimized Miller compensation based on voltage buffer," *Proceedings of the 38th Midwest Symposium on Circuits and Systems*, August 13–16, 1995, pp. 1034–1037.
- Solomon, J. E., "The monolithic op amp: a tutorial study," *IEEE Journal of Solid-State Circuits*, vol. 9, no. 6, December 1974, pp. 314–332.
- Thompson, Marc T., "Design Linear Circuits Using OCTC Calculations," *Electronic Design (Special Analog Issue)*, June 24, 1993, pp. 41–47.
- , "Network Tricks Aid in OCTC," *Electronic Design*, December 16, 1993, pp. 67–70.
- Yang, H. C., and Allstot, D. J., "An equivalent circuit model for two-stage operational amplifiers," *Proceedings of the 1988 IEEE International Symposium on Circuits and Systems*, June 7–9, 1988, pp. 635–638.
- , "Modified modeling of Miller compensation for two-stage operational amplifiers," *Proceedings of the 1991 IEEE International Symposium on Circuits and Systems*, June 11–14, 1991, pp. 2557–2560.

High-Gain Bipolar Amplifiers and BJT Current Mirrors

In This Chapter

- In this chapter, we discuss a more detailed incremental model of the bipolar transistor that takes into account the base-width modulation effect. The resulting resistive elements resulting from base-width modulation have significant design impact on high-gain amplifiers, emitter-followers and current mirrors.

The Need to Augment the Hybrid-Pi Model

The hybrid-pi model considered so far assumes that the incremental output resistance of a bipolar transistor is infinite. However, this is not the case, as anyone can attest to who has looked at transistor curves on a curve tracer. We need to consider the effects of base-width modulation to account for the fact that the output resistance seen at a transistor collector is finite.

A resistively loaded common-emitter amplifier is shown in **Figure 8-1**. To maximize output voltage swing, we'll set the bias point of V_o to $V_{CC}/2$ (through means that are not explicitly shown in this schematic). This, in turn, sets the collector current and hence the transconductance of the transistor:

$$g_m = \frac{|I_C|}{kT/q} = \frac{V_{CC}}{2R_L V_{TH}} \quad [8-1]$$

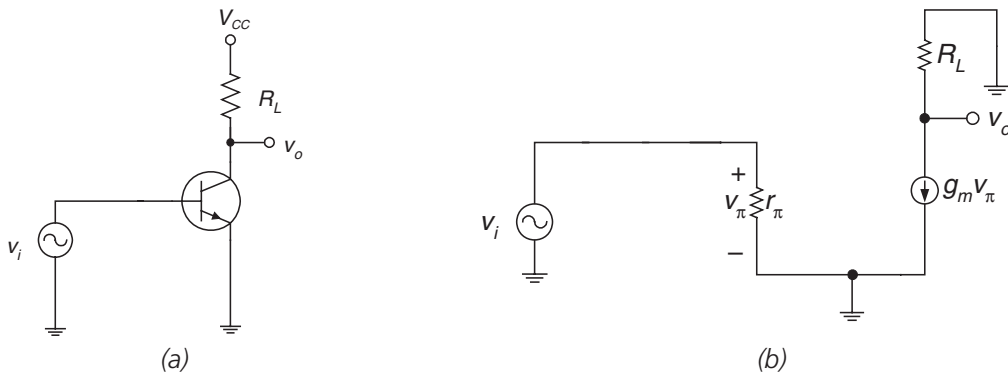


Figure 8-1: Common-emitter amplifier with resistive load. (a) Circuit, omitting biasing details. (b) Low-frequency hybrid-pi model developed so far.

This results in a maximum incremental gain¹ for the resistively loaded common-emitter amplifier as:

$$GAIN_{\max} = -\frac{V_{CC}}{2V_{TH}} \quad [8-2]$$

For $V_{CC} = +12$, this results in a maximum incremental gain of -230 .

In order to get higher gain without arbitrarily high collector voltage, an active load (i.e., a current source) can be used. Let's first consider the ramifications of ignoring base-width modulation and the transistor finite output resistance. Consider the common-emitter amplifier with a current source load (**Figure 8-2a**) where, for simplicity, details of the bias circuit have been omitted. In this case, instead of a collector load resistor we have a collector active load that is the I_{DC} current source.

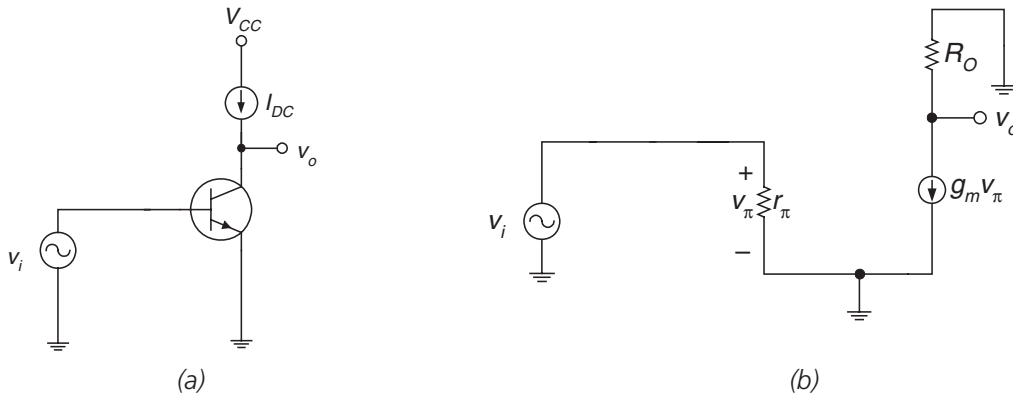


Figure 8-2: Common-emitter amplifier with current source load.
(a) Circuit, omitting biasing details. (b) Incremental model, assuming that the I_{DC} current source has incremental output resistance R_O .

The small-signal model is shown in **Figure 8-2b**. Let's assume that the incremental output resistance of the I_{DC} current source approaches infinity; this means that the load resistor in the incremental model $R_O \Rightarrow \infty$ as well. Note that in this simplified model, the gain is:

$$\frac{v_o}{v_i} = -g_m R_O \Rightarrow \infty \quad [8-3]$$

We know this can't be the case in a real-world circuit. As we'll see later, other transistor internal incremental resistances limit this gain to a finite value.

¹ The incremental gain for the configuration of **Figure 8-1b** is $-g_m R_L$.

Base-Width Modulation

The simple hybrid-pi model used so far leads us to believe that the transistor collector current doesn't change when the collector voltage changes. In other words, the simple model indicates infinite output impedance at a transistor's collector. If you look closely at transistor curves (for instance, on a datasheet or on a curve tracer, as shown in **Figure 8-3a**) you see that there is a finite slope of the collector current in the linear region of operation. This corresponds to the small-signal output resistance of the transistor when operated in the linear operating region.

If we extend a sloped line corresponding to the transistor curves to negative voltages as shown in **Figure 8-3b**, we find that the lines cross the negative voltage axis at the same point V_A . The magnitude of V_A is the *Early voltage*. We can modify the transistor voltage-current relationship to account for the Early voltage as:

$$I_C = I_S \left(e^{\frac{qV_{BE}}{kT}} - 1 \right) \left(1 + \frac{V_{CE}}{V_A} \right) \quad [8-4]$$

This equation captures the functional dependence of collector current on collector-emitter voltage.

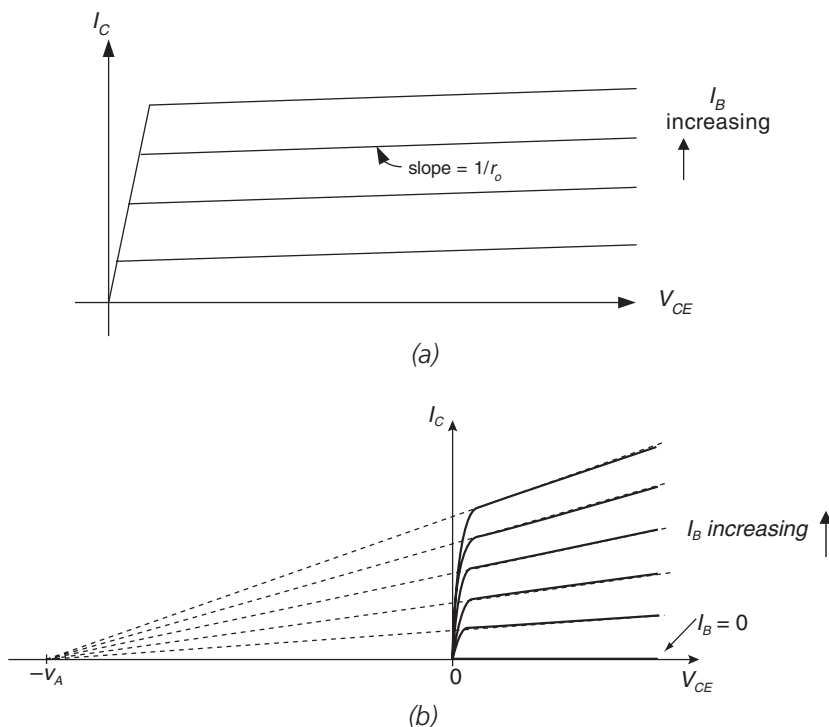


Figure 8-3: NPN transistor generic V-I curve. (a) Curves showing effects of finite transistor output resistance r_o . (b) Definition of Early voltage V_A .

This effect is due to the widening/narrowing of the collector-base depletion region when V_{CB} changes (**Figure 8-4**), which in turn changes the effective base width. When V_{CB} increases, the width of the collector-base depletion region increases from W to $W + \Delta W$. Since the collector current is proportional to the *slope* of the minority carrier concentration in the base (in this case, electrons) the collector current changes when V_{CB} changes. This process is called *base-width modulation* and was first described by James Early² in 1952.

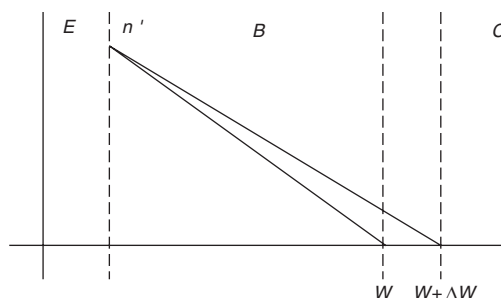


Figure 8-4: Illustration of base-width modulation. The slope of the minority carrier concentration n' varies as V_{CE} varies, resulting in a change in the collector current.

Base-width modulation causes both the collector current and base current of the transistor to change when the collector voltage changes. This small-signal effect means that a transistor current source doesn't have infinite output impedance. These effects are modeled by adding two more resistors to the hybrid-pi model (**Figure 8-5**). This model should be used when analyzing transistor stages with gains in excess of a few hundred.

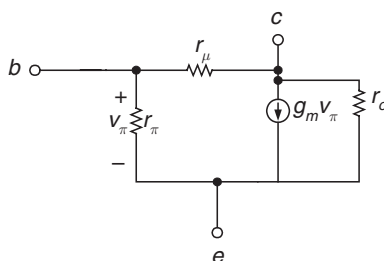


Figure 8-5: Transistor extended hybrid-pi model (low frequency) including effects of base-width modulation, resulting in added circuit elements r_o and $r_μ$. Resistance r_o models the fact that collector current varies as V_{CE} varies. Resistance $r_μ$ models the effects of the extra base current needed to support this current.

² See the James Early reference at the end of this chapter.

Finding Parameters from a Transistor Datasheet

A detailed analysis shows that the output resistance of the transistor r_o is inversely proportional to collector current,³ or:

$$r_o = \frac{1}{\eta g_m} = \frac{V_A}{\left(\frac{kT}{q}\right) g_m} \quad [8-5]$$

where η is a constant called the *base-width modulation factor* that has typical values of 10^{-3} to 10^{-4} . Changes in collector-base voltage have a much smaller effect on output current than changes in base-emitter voltage. Therefore, the effects of base-width modulation and output resistance r_o are only significant if the gain of the amplifier approaches $1/\eta$. A rule-of-thumb is that we need to consider r_o in gain calculations when the gain of our amplifier is greater than a few hundred.

The change in collector current also results in a proportional change of base current, with the proportionality constant being the small-signal current gain h_{fe} , resulting in:

$$r_\mu = h_{fe} r_o = \frac{h_{fe}}{\eta g_m} \quad [8-6]$$

Transistor manufacturers sometimes specify transistors using two-port h -parameters, corresponding to the circuit of **Figure 8-6**.

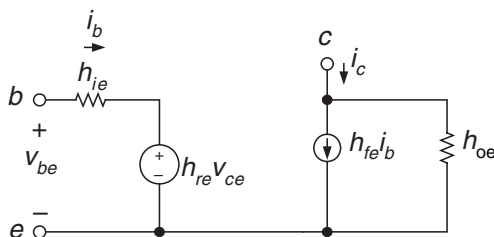


Figure 8-6: Transistor model showing two-port h -parameters.

The equations describing the operation of this circuit are as follows:

$$\begin{aligned} v_{be} &= i_b h_{ie} + h_{re} v_{ce} \\ i_c &= h_{fe} i_b + h_{oe} v_{ce} \end{aligned} \quad [8-7]$$

Comparing this to the extended hybrid-pi model results in:

$$\begin{aligned} h_{ie} &= r_\pi \\ h_{re} &= \eta \\ h_{fe} &= \beta_o \\ h_{oe} &= \eta g_m \end{aligned} \quad [8-8]$$

³ For a detailed mathematical derivation, see, e.g., Gray, Hurst, Lewis and Meyer, pp. 14–16 or P. E. Gray, et al., *Physical Electronics and Circuit Models of Transistors*, SEEC volume 2, pp. 149–152.

Shown below are the h -parameters from the 2N3904 transistor⁴ datasheet (**Figure 8-7**).

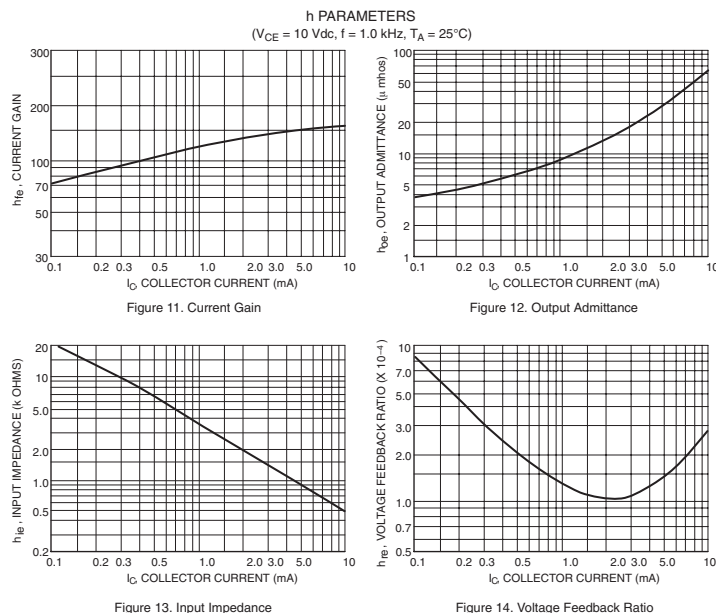


Figure 8-7: h -parameters from 2N3904 datasheet.

Common-Emitter Amplifier with Current Source Load

Let's assume that we have a common-emitter amplifier biased with an ideal current source of 1 milliamp (**Figure 8-8a**). Our assumption is that the current source is ideal—hence the current source has an infinite incremental output resistance. Let's find the gain of the amplifier, given that the base width modulation factor of the transistor $\eta = 3.3 \times 10^{-4}$.

A small-signal model of this circuit is shown in **Figure 8-8b**, where we have used the transistor extended hybrid- π model.

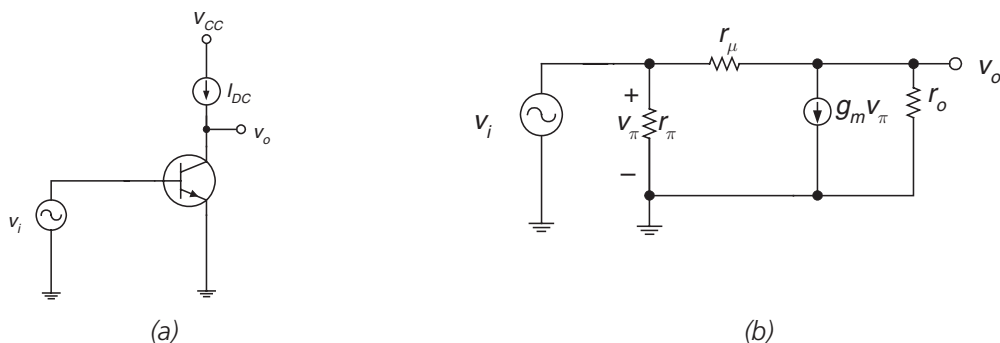


Figure 8-8: Common-emitter with active load. (a) Circuit. (b) Small-signal model using extended hybrid- π model for the transistor.

⁴ Found at www.onsemi.com, reprinted with permission of On Semiconductor.

Applying KCL at the output (v_o) node results in:⁵

$$(v_i - v_o)g_\mu - v_o g_o - v_i g_m = 0 \quad [8-9]$$

Solving for gain v_o/v_i results in:⁶

$$\frac{v_o}{v_i} = \frac{g_\mu - g_m}{g_o + g_\mu} \approx \frac{-g_m}{g_o} \approx \frac{-1}{\eta} \quad [8-10]$$

For this example, we predict a gain of -3000 . Now, achieving a gain as high as this is contingent on how *ideal* a current source we can build. We'll see in the next section methods for building a current source with a high output resistance.

Building Blocks

Incremental output resistance of bipolar current source

Let's attempt to answer the question of how ideal a current source built with bipolar junction transistors can be. **Figure 8-9a** shows a current source with the transistor having an emitter resistor R_E . The small-signal model is shown in **Figure 8-9b**, where we've added a test current source i_t in order to calculate the incremental output resistance of the current source, as seen at the transistor collector.

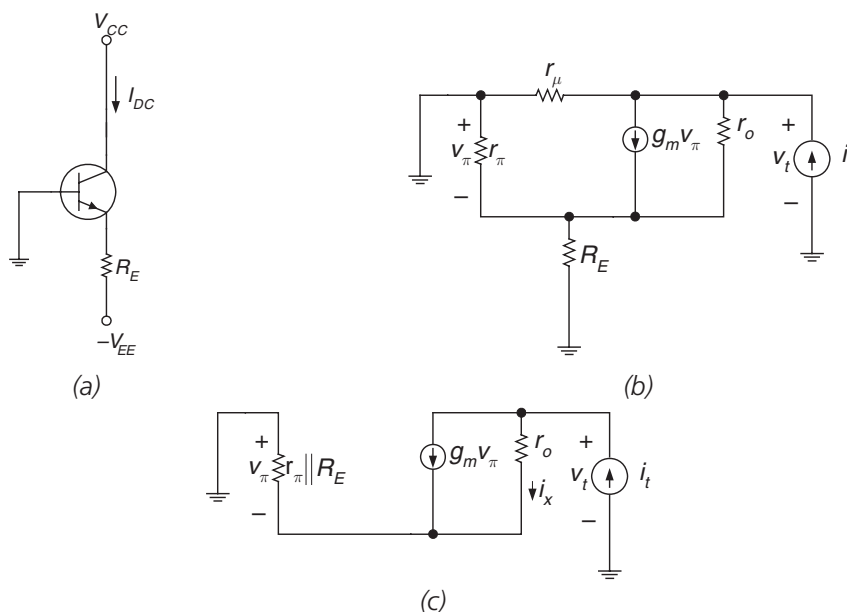


Figure 8-9: Current source with finite emitter resistor R_E .
 (a) Schematic. (b) Small-signal model. (c) Simplified small-signal model.

⁵ Again, we note that the math is a little bit easier if we use conductances instead of resistances. For instance, in this circuit, $g_\mu = 1/r_\mu$.

⁶ Note that $g_m \gg g_\mu$ and that $g_o \gg g_\mu$.

The analysis of this circuit is greatly simplified by combining r_π and R_E into a common resistor, and by removing r_μ (and adding it in parallel later). Using **Figure 8-9c**, the test voltage v_t is given by:

$$v_t = i_t [r_\pi \parallel R_E] + i_x r_o \quad [8-11]$$

The current i_x is the difference between i_t and $g_m v_\pi$, or:

$$i_x = i_t - g_m v_\pi \quad [8-12]$$

We also know that v_π is related to i_t by the parallel combination of r_π and R_E . So, we can solve for v_t as.

$$v_t = i_t [r_\pi \parallel R_E] + [i_t + g_m i_t [r_\pi \parallel R_E]] r_o \quad [8-13]$$

Therefore, the output resistance of the current source is:

$$r_{out} = \frac{v_t}{i_t} = [r_\pi \parallel R_E] + [1 + g_m [r_\pi \parallel R_E]] r_o \approx \frac{[1 + g_m [r_\pi \parallel R_E]]}{g_o} \quad [8-14]$$

For the final result, we have to add r_μ back in parallel, resulting in:

$$r_{out} \approx r_\mu \parallel \frac{[1 + g_m [r_\pi \parallel R_E]]}{g_o} \quad [8-15]$$

There are several important limiting cases. If the emitter is grounded (i.e., if $R_E = 0$) the output resistance is:

$$r_{out} \approx r_o \text{ for } R_E = 0 \quad [8-16]$$

If the emitter resistance is large compared to r_π , the output resistance is:

$$r_{out} \approx \frac{r_\mu}{2} \text{ for } R_E \gg r_\pi \quad [8-17]$$

The lesson here is that we can build a current source with a very high output resistance provided that we have an emitter resistance that is large compared to r_π . One way to accomplish this without using large-valued resistors is to use a cascode current source (**Figure 8-10**).

Transistor Q_2 is the output transistor, which carries the desired output current I_{DC} . Transistor Q_1 provides a high output impedance seen by the emitter of Q_2 . This ensures that the output resistance at the collector of Q_2 is high. The inclusion of R_{E1} increases the output resistance of the collector of Q_1 .

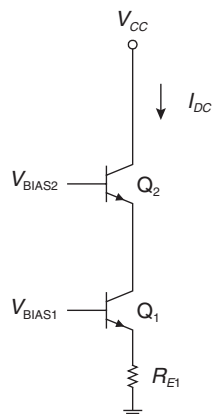


Figure 8-10: Cascode current source. The output resistance at the collector of Q_2 is increased by inclusion of Q_1 in its emitter.

Emitter-follower incremental input resistance

The incremental input resistance of an emitter-follower (**Figure 8-11a**) can be calculated using the circuit of **Figure 8-11b**. For instance, if you have a common-emitter gain stage followed by an emitter-follower buffer, you would like to know how much the emitter-follower loads the gain stage. In **Figure 8-11b** is the incremental circuit of the emitter-follower, including extended hybrid-pi model values r_o and r_μ .

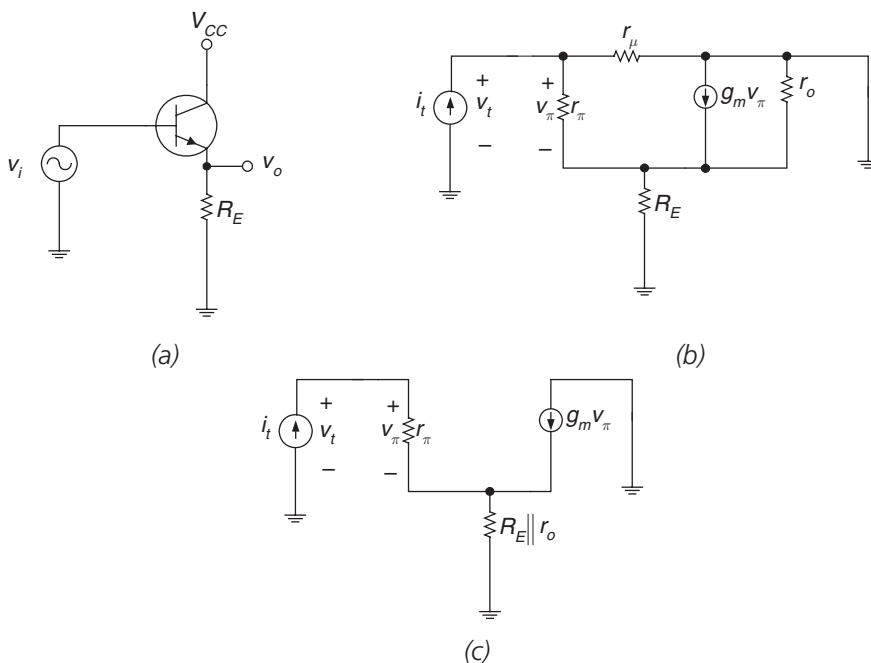


Figure 8-11: Emitter-follower (a) Circuit. (b) Incremental model using extended hybrid-pi transistor model. (c) Circuit simplified by removal of r_μ .

Note that in **Figure 8-11b**, incrementally, the transistor output resistance appears in parallel with R_E . Hence, the output resistance is (with r_μ taken out to simplify the math as in **Figure 8-11c**) is:

$$r_{in} = r_\pi + (1 + h_{fe})[R_E \parallel r_o] \quad [8-18]$$

Now, let's put back in r_μ in parallel, to arrive at the final result:

$$r_{in} = r_\mu \parallel \left(r_\pi + (1 + h_{fe})[R_E \parallel r_o] \right) \quad [8-19]$$

Now, the limit for very big emitter resistance R_E is:

$$r_{in} \approx \frac{r_\mu}{2} \text{ if } R_E \gg r_o \quad [8-20]$$

The limit for very small emitter resistance is:

$$r_{in} \approx r_\pi \text{ if } R_E \rightarrow 0 \quad [8-21]$$

For intermediate values of R_E we find:

$$r_{in} \approx h_{fe} R_E \text{ if } r_\pi \ll R_E \ll r_o \quad [8-22]$$

Example 8.1: Incremental input resistance of emitter-follower

Find the incremental input resistance of the emitter-follower of **Figure 8-12** with the following transistor parameters and bias conditions: collector current $I_C = 1$ milliamp; emitter resistor $R_E = 10 \text{ k}\Omega$; small-signal current gain $h_{fe} = 100$; and base-width modulation factor $\eta = 10^{-4}$.

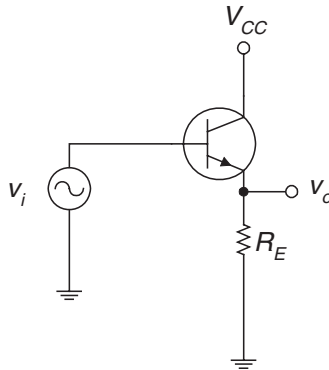


Figure 8-12: Emitter-follower.

SOLUTION

We find the small signal parameters as follows:

$$\begin{aligned}
 g_m &= \frac{I_C}{V_{TH}} = \frac{10^{-3}}{0.026} = 0.038 \text{ A/V} \\
 r_\pi &= \frac{h_{fe}}{g_m} = \frac{100}{0.038} = 2600\Omega \\
 r_o &= \frac{1}{\eta g_m} = \frac{1}{(10^{-4})(0.038)} = 260 \text{ k}\Omega \\
 r_\mu &= h_{fe} r_o = 26 \text{ M}\Omega
 \end{aligned} \tag{8-23}$$

The small-signal model for this emitter-follower is shown in **Figure 8-13**.

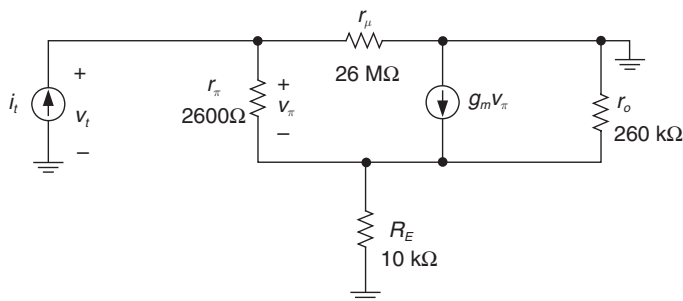


Figure 8-13: Emitter-follower small-signal model.

The incremental input resistance is:

$$r_{in} = r_\mu \parallel \left(r_\pi + (1 + h_{fe}) [R_E \parallel r_o] \right) \approx (26 \text{ M}\Omega) \parallel \left((100) [(10\text{k}) \parallel 260\text{k}] \right) \approx 1 \text{ M}\Omega \tag{8-24}$$

Note that this input resistance is approximately $r_{in} \approx h_{fe} R_E$, which is the result we expect for intermediate values of emitter resistance.

Current mirrors

Current mirrors, also called *current repeaters*, are commonly used as biasing elements and as active loads in amplifiers.⁷ A basic bipolar junction transistor current mirror is made up of a first diode-connected transistor⁸ through which flows a control current, and a second output transistor where the output current is controlled by the voltage generated across the first diode-connected transistor. Hence, in a current mirror, the base-to-emitter voltages of the two transistors are identical, causing currents in the two transistors to be proportional.⁹ Current mirrors can be constructed from NPN transistors, PNP transistors, or combinations of the two. Some basic current mirror circuits and the equivalent circuits (with the diode-connected transistor replaced by a simple diode in the schematic) are shown in **Figure 8-14**.

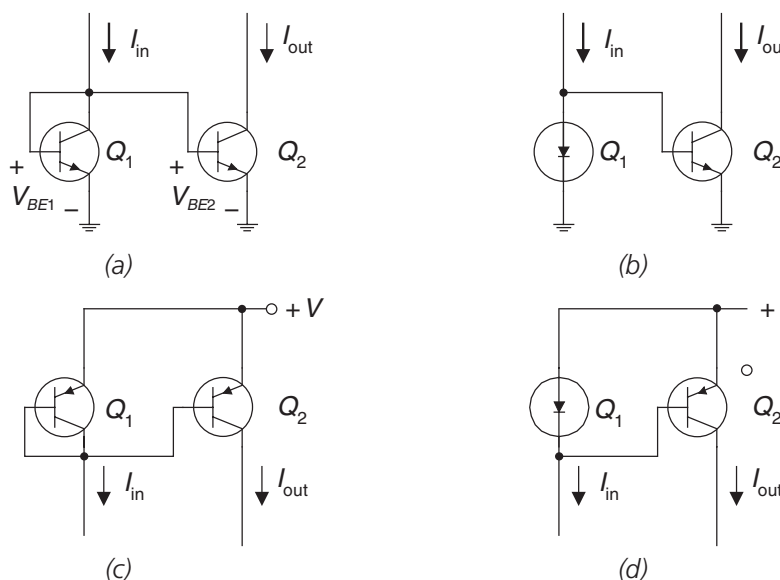


Figure 8-14: Basic current mirrors and equivalent circuits. (a) Basic NPN mirror, showing two transistors with equal V_{BE} s. (b) Basic NPN mirror equivalent circuit with diode connected transistor Q_1 replaced by a diode for illustrative purposes. (c) Basic PNP mirror. (d) Basic PNP mirror equivalent circuit.

⁷ See, for instance, J. Roberge, *Operational Amplifiers Theory and Practice*, p. 393; Gray, Hurst Lewis and Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th edition, p. 255.

⁸ A “diode connected” bipolar junction transistor is one where the collector is shorted to the base and the transistor is operated as a diode utilizing the base-emitter junction. See J. Roberge, *Operational Amplifiers: Theory and Practice*, p. 390.

⁹ This analysis assumes that base currents are negligible and assumes that the transistor has high output impedance. For the collector currents of the two transistors to be equal, the transistors must be matched and have the same emitter areas. If the emitter areas are made differently, there is a scale factor relating the input and output currents.

Current mirrors are found in many integrated circuit operational amplifiers, including the Fairchild $\mu A709$, $\mu A741$, and $\mu A776$, Analog Devices's OP07, National Semiconductor's LM301A and LM308, among others. Current mirrors can be fabricated with either bipolar junction transistors (BJTs) or MOSFETs.

The basic bipolar current mirror has an input/output gain error due to the finite β_F of the transistor. In **Figure 8-15**, we see that the two transistors are operating at the same V_{BE} , and hence to first order the transistors have the same collector current. However, the input current is not the collector current, because I_{in} also supplies base currents to Q_1 and Q_2 . We can express this relationship as:

$$I_{in} = I_{C1} + \frac{I_{C1}}{\beta_F} + \frac{I_{out}}{\beta_F} \quad [8-25]$$

Solving for current mirror gain results in:

$$\frac{I_{out}}{I_{in}} = \frac{1}{1 + \frac{2}{\beta_F}} \quad [8-26]$$

We see that the error is inversely proportional to the value of DC current gain β_F . For current gain $\beta_F = 100$, this results in approximately 2% error in output current.

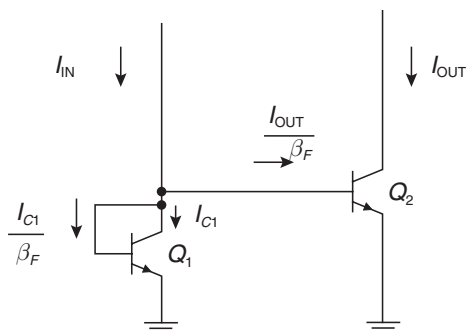


Figure 8-15: Circuit for determining gain error of basic current mirror. Intermediate circuit currents are shown for clarity.

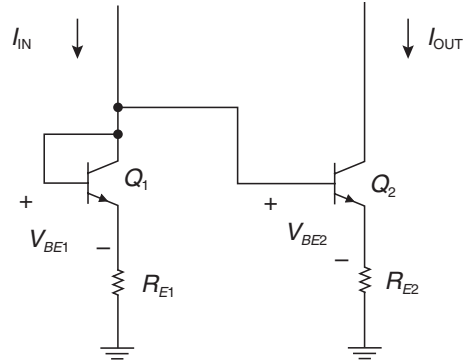
Basic current mirror with emitter degeneration

One way to reduce current mirror errors due to mismatched V_{BE} s is to use some emitter degeneration in the current mirror (**Figure 8-16**). We see that an equation relating the two emitter currents of Q_1 and Q_2 is:

$$I_{E1}R_{E1} + V_{BE1} = I_{E2}R_{E2} + V_{BE2} \quad [8-27]$$

We see that if $I_E R_{E1} \gg V_{BE1}$ and $I_E R_{E2} \gg V_{BE2}$ any effects of V_{BE} mismatches are reduced, assuming the emitter resistors are well matched. Inclusion of the emitter resistors also increases the output of the current source. We can also adjust the mirror ratio by adjusting the relative ratio of R_1 to R_2 .

Figure 8-16: Basic current mirror with emitter degeneration resistor R_{E1} and R_{E2} .



Current mirror with “beta helper”

The dependence of output current with transistor β_F can be improved by adding another transistor and altering the configuration of the current mirror as in **Figure 8-17**. The values of each current are as shown; transistor Q_3 provides base current to both Q_1 and Q_2 and hence reduces the dependence of the output current on transistor current gain.

In analyzing this circuit, we recognize that Q_1 and Q_2 operate the same V_{BE} and hence their collector currents are the same (I_{OUT}). Transistor Q_3 provides base current for Q_1 and Q_2 , and hence its emitter current is $2I_{OUT}/\beta_F$. The input current (I_{IN}) is not exactly the same as the collector current of Q_1 , since we have the base current of Q_3 as well. We find I_{OUT} as follows:

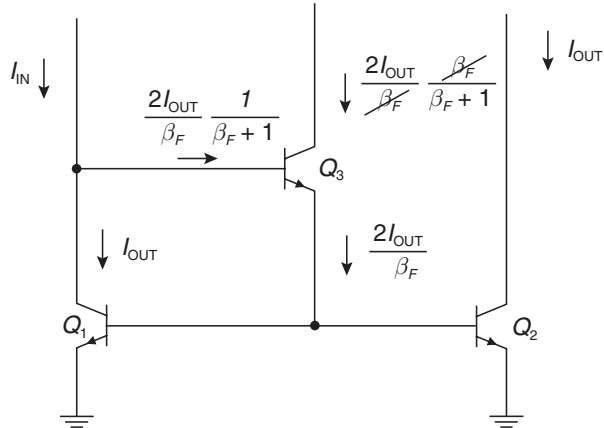
$$I_{IN} = I_{OUT} + \left(\frac{2I_{OUT}}{\beta_F} \right) \left(\frac{1}{\beta_F + 1} \right) \quad [8-28]$$

Through some algebraic manipulation we find that the output and input currents are related as:

$$\frac{I_{OUT}}{I_{IN}} = \frac{\beta_F^2 + \beta_F}{\beta_F^2 + \beta_F + 2} = \frac{1}{1 + \frac{2}{\beta_F^2 + \beta_F}} \quad [8-29]$$

We see that the error is inversely proportional to the square of β_F , so the error is reduced as compared to the simple current mirror.

Figure 8-17: Current mirror with beta helper, showing intermediate currents to help in calculating the output/input ratio.



Wilson current mirror

The Wilson current mirror (**Figure 8-18**) reduces errors due to finite β_F by the use of negative feedback. Let's assume that I_{B3} increases a little bit. This tends to increase the collector current of Q_3 and hence also the collector current Q_2 . This increases the collector current of Q_2 , and tends to reduce I_{B3} (since an increase in I_{C1} steals current away from the base of Q_3). This is negative feedback. Shown in **Figure 8-18** are intermediate currents which help us determine the input-output relationship.

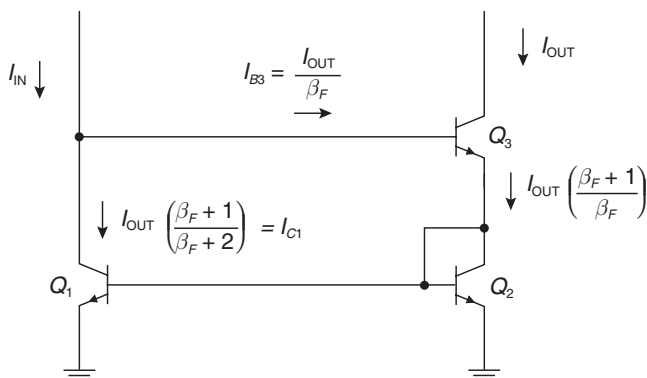


Figure 8-18: Wilson current mirror.

A detailed analysis similar to that we did with the mirror with beta helper results in the input/output current relationship for the Wilson current mirror as:

$$I_{IN} = \frac{I_{OUT}}{\beta_F} + I_{OUT} \left(\frac{\beta_F + 1}{\beta_F + 2} \right) \quad [8-30]$$

After some algebra, we find for the Wilson mirror the transfer function relating input and output current as:

$$\frac{I_{OUT}}{I_{IN}} = \frac{\beta_F^2 + 2\beta_F}{\beta_F^2 + 2\beta_F + 2} = \frac{1}{1 + \frac{2}{\beta_F^2 + 2\beta_F}} \quad [8-31]$$

Cascode current mirror

A cascode current mirror (**Figure 8-19**) has an output stack of transistors. We have seen earlier in our work on the extended hybrid-pi model that the output resistance of a transistor with emitter resistance $R_E \gg r_\pi$ results in an output resistance approaching $r_o/2$. One way to accomplish this is to stack output transistors. This cascode current mirror has a very high output impedance, since output transistor Q_4 has a large resistance seen at its emitter, due to the output resistance of transistor Q_2 .

Widlar current mirror

A Widlar mirror can be used when low levels of output currents are needed. In the forward active region, the base-emitter voltage of a transistor is given by:

$$V_{BE} \approx \frac{kT}{q} \ln \left(\frac{I_C}{I_S} \right) \quad [8-32]$$

where kT/q is the thermal voltage and I_S is the reverse saturation current of the transistor. In **Figure 8-20**, we can solve KVL around the Q_1 and Q_2 base-emitter loops, resulting in:

$$V_{BE1} = V_{BE2} + I_{E2} R_E \quad [8-33]$$

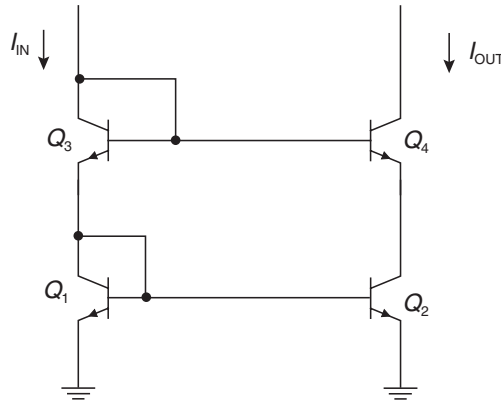


Figure 8-19: Cascode current mirror.

This means that the emitter resistor R_E steals away some of the voltage available to drive the base-emitter junction of Q_2 . Hence Q_2 runs at a lower collector current than Q_1 , but Q_1 still maintains some level of control.

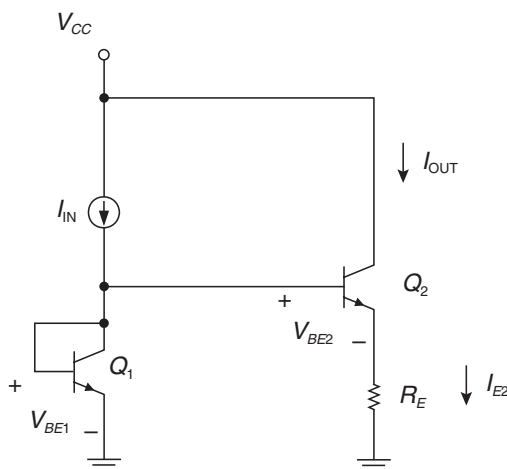


Figure 8-20: Widlar current mirror.

Example 8.2: Widlar current mirror

For the Widlar current mirror of **Figure 8-20**, find the collector current of transistor Q_2 for an input current $I_{in} = 100$ microamps and emitter resistor $R_E = 10$ k Ω . Assume that each transistor has a reverse saturation current $I_S = 10^{-6}$ A, that $\beta_F \gg 1$ for each transistor, and that the transistors operate at a temperature of 300K.

SOLUTION

Since we assume that $\beta_F \gg 1$, then $I_E \approx I_C$ and we'll ignore base currents.

$$V_{BE1} = V_{BE2} + I_{E2}R_E \Rightarrow \frac{kT}{q} \ln \left(\frac{I_{IN}}{I_{S1}} \right) \approx \frac{kT}{q} \ln \left(\frac{I_{OUT}}{I_{S2}} \right) + I_{OUT}R_E \quad [8-34]$$

Let's simplify this expression by combining the logarithmic terms:

$$\frac{kT}{q} \ln \left(\frac{I_{IN}}{I_{OUT}} \right) \approx I_{OUT}R_E \quad [8-35]$$

Note that the dependence on the reverse saturation current drops out. Iteratively solving this equation with $I_{IN} = 100$ μ A, we find that $I_{OUT} \approx 6.9$ μ A.

Example 8.3: Widlar mirror incremental output resistance

For the Widlar current mirror of **Figure 8-20**, find the incremental output resistance measured at the collector of Q_2 . Again, assume that $R_E = 10 \text{ k}\Omega$. For Q_2 , assume the following parameters: $h_{fe} = 100$; $\eta = 3.3 \times 10^{-4}$.

SOLUTION

We find the small-signal parameters as follows:

$$\begin{aligned} g_m &= \frac{I_{C2}}{V_{TH}} = \frac{6.9 \times 10^{-6}}{0.026} = 2.65 \times 10^{-4} \text{ A/V} \\ r_\pi &= \frac{h_{fe}}{g_m} = \frac{100}{2.65 \times 10^{-4}} = 377 \text{ k}\Omega \\ r_o &= \frac{1}{\eta g_m} = \frac{1}{(3.3 \times 10^{-4})(2.65 \times 10^{-4})} = 11.4 \text{ M}\Omega \\ r_\mu &= h_{fe} r_o = 1140 \text{ M}\Omega \end{aligned} \quad [8-36]$$

The full expression for the output resistance of transistor Q_2 with emitter degeneration is:

$$r_{out} \approx r_\mu \left\| \frac{[1 + g_m [r_\pi \parallel R_E]]}{g_o} \right\| \quad [8-37]$$

Note that in this case $R_E \ll r_\pi$ and hence we can approximate the output resistance as:

$$r_{out} \approx g_m R_E r_o \approx (2.65 \times 10^{-4})(10,000)(11.4 \text{ M}\Omega) \approx 30.2 \text{ M}\Omega \quad [8-38]$$

Example 8.4: Current mirror error due to mismatched VCEs

A circuit illustrating the cause of current mirror mismatch due to the Early effect is shown in **Figure 8-21**. Transistor Q_1 is an input diode-connected transistor biased by a 10-milliamp current source (I_I). The SPICE plot¹⁰ of **Figure 8-21b** shows the collector current of transistor Q_2 as the power supply voltage is varied from 0 to 20 volts. In this case, the V_{CE} of Q_1 is fixed at a V_{BE} drop (due to the diode connection) while the V_{CE} of Q_2 is significantly higher. The collector current of Q_2 varies significantly over the range of power supply voltages. The collector current of Q_2 rises to ~10 milliamps when the supply voltage reaches ~0.5 volts. This is when transistor Q_2 leaves saturation and enters the forward-active region.

¹⁰ The schematic of **Figure 8-21a** shows the node voltages for a power supply voltage of +12V.

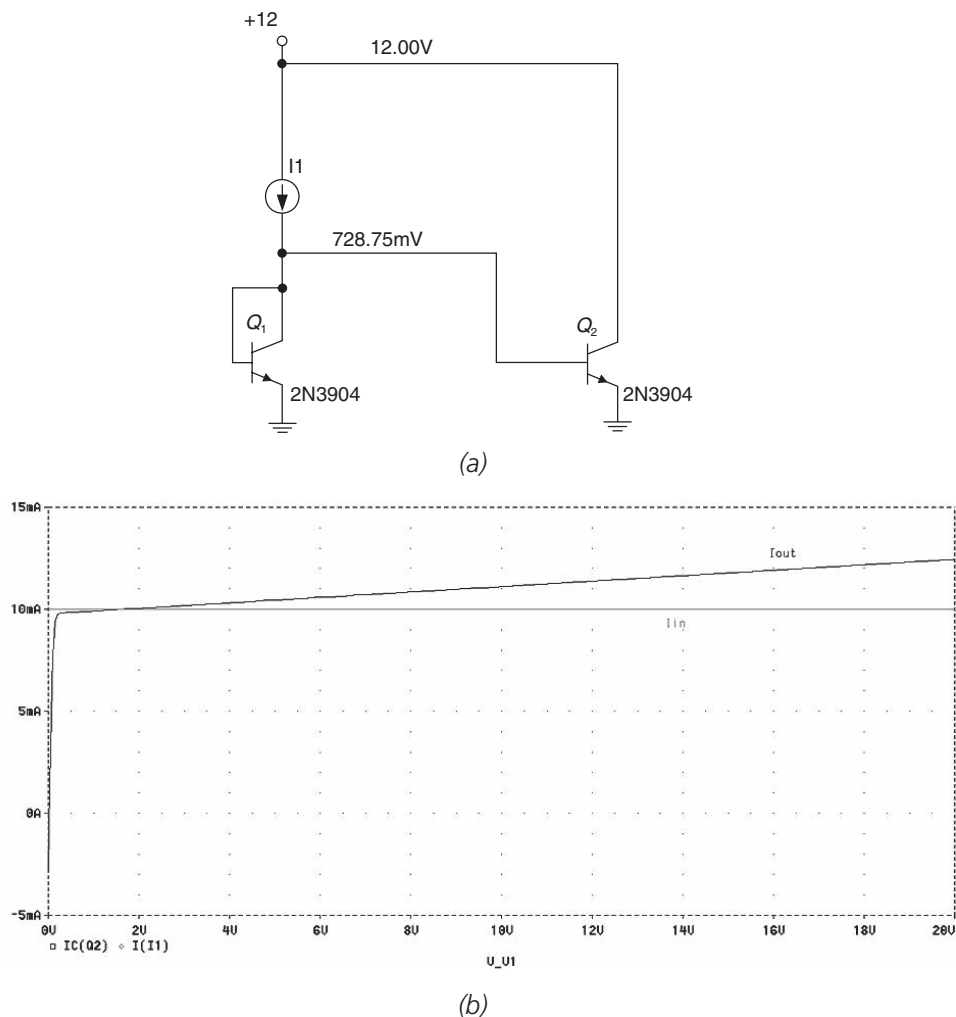


Figure 8-21: Circuit illustrating current mirror error due to mismatched V_{CE} s. (a) Circuit. (b) PSPICE simulation result showing variation in I_{out} as power supply voltage varies from 0 to 20 volts.

Example 8.5: Design example—high-gain amplifier

Consider the high-gain transistor amplifier in **Figure 8-22**. We will find the gain of this amplifier, given that the base-width modulation factors for the NPN and the PNP transistors are $\eta_{npn} = 3.2 \times 10^{-4}$ and $\eta_{pnp} = 1.2 \times 10^{-3}$. Other parameters of interest are $h_{fe,npn} = 203$ and $h_{fe,pnp} = 164$.

Transistor Q_1 is a standard common-emitter stage, biased with the PNP current source Q_2 . The 1000-k Ω resistor feedback biases the amplifier so that both Q_1 and Q_2 are biased in their linear operating region. Note that we have made the emitter resistor of Q_2 (R_{E2}) large enough so that the incremental output resistance r_{out2} of Q_2 is large. This emitter resistor also sets a collector current of approximately 1 milliamp in Q_1 and Q_2 .

In **Figure 8-22b**, we draw the incremental model, replacing transistor Q_2 by its equivalent output resistance r_{out2} . Parameters for Q_2 are as follows:

$$\begin{aligned}
 I_{C2} &\approx 1 \text{ mA} \\
 g_{m2} &= \frac{1 \text{ mA}}{26 \text{ mV}} \approx 0.038 \text{ A/V} \\
 r_{\pi2} &= \frac{h_{fe2}}{g_{m2}} = \frac{164}{0.038} = 4316 \Omega \\
 r_{o2} &= \frac{1}{\eta_{pnp} g_{m2}} = \frac{1}{(1.2 \times 10^{-3})(0.038)} = 2.2 \times 10^4 \Omega \\
 r_{\mu2} &= h_{fe2} r_{o2} = (164)(2.2 \times 10^4) = 3.6 \times 10^6 \Omega \\
 r_{out2} &= r_{\mu2} \parallel r_{o2} \left[1 + g_{m2} [r_{\pi2} \parallel R_E] \right] = 5.9 \times 10^5 \Omega
 \end{aligned} \tag{8-39}$$

For transistor Q_1 , bias level and incremental parameters are:

$$\begin{aligned}
 I_{C1} &\approx 1 \text{ mA} \\
 g_{m1} &= \frac{1 \text{ mA}}{26 \text{ mV}} \approx 0.038 \text{ A/V} \\
 r_{o1} &= \frac{1}{\eta_{npn} g_{m1}} = \frac{1}{(3.2 \times 10^{-4})(0.038)} = 8.2 \times 10^4 \Omega \\
 r_{\mu1} &= h_{fe1} r_{o1} = (203)(8.2 \times 10^4) = 1.7 \times 10^7 \Omega
 \end{aligned} \tag{8-40}$$

The node equation for finding the gain is:

$$\begin{aligned}
 (v_i - v_o)G'_f - g_{m1}v_i - v_o G_o &= 0 \\
 G'_f &= \frac{1}{R_f \parallel r_{\mu1}} \\
 G_o &= \frac{1}{r_{o1} \parallel r_{out2}}
 \end{aligned} \tag{8-41}$$

Reducing this results in:

$$\begin{aligned}
 v_i (G'_f - g_{m1}) &= v_o (G'_f + G_o) \\
 \frac{v_o}{v_i} &= \frac{(G'_f - g_{m1})}{(G'_f + G_o)} \approx \frac{-g_{m1}}{(G'_f + G_o)} \approx -g_{m1} (R_f \parallel r_{\mu1} \parallel r_{o1} \parallel r_{out2}) = -2541
 \end{aligned} \tag{8-42}$$

A PSPICE simulation (**Figure 8-23**) shows that the midband gain is approximately as calculated. The low-frequency rolloff can be adjusted by modifying the value of the coupling capacitor C_c .

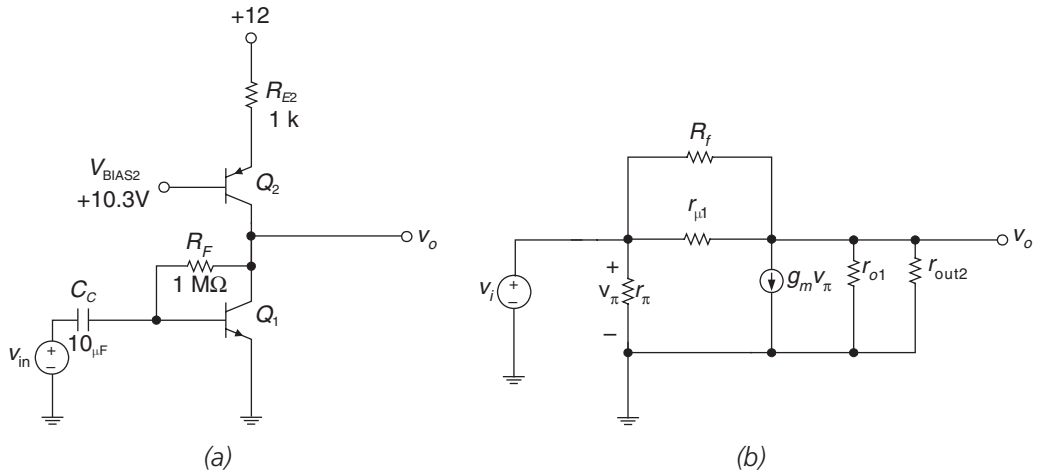


Figure 8-22: High-gain amplifier. (a) Circuit. (b) Small-signal model.

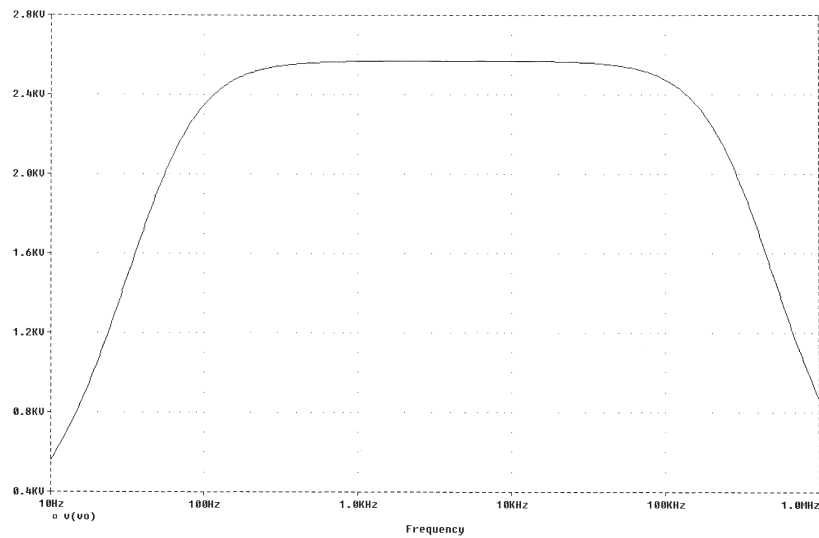


Figure 8-23: High-gain amplifier—PSPICE simulation result, showing a midband gain of approximately 2500.

Example 8.6: Another high-gain amplifier example

Let's design an AC-coupled transistor amplifier with gain magnitude of $|A_v| > 1000$, using high-gain techniques. We're driving a 100-picofarad capacitive load with the amplifier, so we need an emitter-follower at the output to isolate the high-gain node from the capacitive load.

Assume that you have at your disposal transistors with $\eta_{npn} = 6.7 \times 10^{-4}$, $\eta_{pnp} = 1.8 \times 10^{-4}$, $C_\mu = 2$ picofarads, $f_T = 300$ MHz, $h_{fe,npn} = 200$ and $h_{fe,pnp} = 175$. In this example we won't worry much about biasing details; i.e., you can assume that your collector currents will magically run at the correct bias levels. We'll find a design that meets the gain specification, then estimate the high- and low-frequency breakpoints using open-circuit and short-circuit time constants.

SOLUTION

An initial iteration on a circuit topology is shown in **Figure 8-24a**. Transistor Q_1 is the common-emitter amplifier loaded by current source Q_2 . The base bias voltage V_{BIAS2} at the base of Q_2 and the emitter resistor R_{E2} set the quiescent current of Q_1 and Q_2 . Feedback resistor R_F biases Q_1 in the forward active region. We'll assume that a value of 1 M Ω suffices for R_F , but the actual value depends on the DC current gain of transistor Q_1 and the desired output voltage at the collector of Q_1 . Transistor Q_3 buffers the load capacitor from the high-gain node found at the collector of Q_1 . In order to achieve high gain, we want current source Q_2 to have a high output resistance and we want emitter-follower Q_3 to have a high input resistance. As a sanity check, we know that the best gain that we can achieve with this topology is:

$$|A_v| < \frac{1}{\eta_{npn}} < \frac{1}{6.7 \times 10^{-4}} < 1492 \quad [8-43]$$

Shown in **Figure 8-24b** is the small-signal model¹¹ of this amplifier where we can find the input-output relationship between the input v_{in} and the incremental voltage v_{c1} at the collector of Q_1 . For this incremental analysis we have replaced the current source Q_2 by a resistance r_{out2} which is the output resistance of the current source. Likewise, we've replaced emitter-follower Q_3 by a resistance r_{in3} which is the input resistance of the emitter-follower. We'll assume that the voltage gain from v_{c1} to v_o is ~ 1 , an assumption that needs to be checked later on.

To simplify the gain analysis, we'll redraw the incremental circuit, resulting in **Figure 8-24c**. We have lumped the parallel combination of r_{o1} , r_{out2} and r_{in3} into a single resistor which we'll call R'_L . We have lumped the parallel combination of $r_{\mu 1}$ and R_F into a feedback resistance R'_F . Mathematically, we express this as:

$$\begin{aligned} R'_L &= r_{o1} \parallel r_{out2} \parallel r_{in3} \\ R'_F &= R_F \parallel r_{\mu 1} \end{aligned} \quad [8-44]$$

If we apply KCL at the output node v_{c1} we find:

$$(v_{in} - v_{c1})G'_F - g_{m1}v_{in} - v_{c1}G'_L = 0 \quad [8-45]$$

¹¹ In this analysis, we assume that $r_x = 0$ for each transistor.

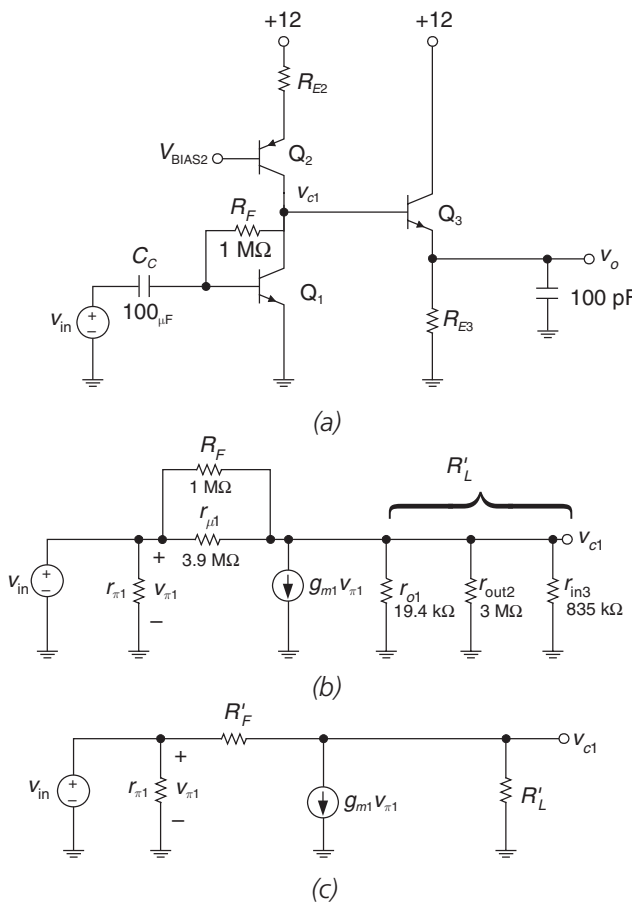
Next, solving for v_{c1}/v_{in} we find:

$$\frac{v_{c1}}{v_{in}} = \frac{G'_F - g_{m1}}{G'_F + G'_L} \approx -\frac{g_{m1}}{G'_F + G'_L} \approx -\frac{g_{m1}}{G_F + g_{\mu 1} + g_{o1} + g_{out2} + g_{in3}} \quad [8-46]$$

Let's somewhat arbitrarily set the collector currents of Q_1 and Q_2 to be 2 milliamps. We find the following small-signal parameters for common-emitter amplifier transistor Q_1 :

$$\begin{aligned} g_{m1} &= \frac{I_{C1}}{\left(\frac{kT}{q}\right)} = \frac{0.002}{0.026} = 0.077 \text{ A/V} \\ r_{\pi 1} &= \frac{h_{fe1}}{g_{m1}} = \frac{200}{0.077} = 2.6 \text{ k}\Omega \\ r_{o1} &= \frac{1}{\eta_{npn} g_{m1}} = \frac{1}{(6.7 \times 10^{-4})(0.077)} = 19.4 \text{ k}\Omega \\ r_{\mu 1} &= h_{fe1} r_{o1} = (200)(19.4 \text{ k}\Omega) = 3.9 \text{ M}\Omega \end{aligned} \quad [8-47]$$

Figure 8-24: High-gain amplifier design example.
(a) Circuit. (b) Small-signal model. (c) Simplified small-signal model.



Chapter 8

We now see that the value of r_{o1} is much smaller than R_F and $r_{\mu1}$, so we can approximate our gain relationship further as:

$$\frac{v_{c1}}{v_{in}} \approx - \frac{g_{m1}}{g_{o1} + g_{out2} + g_{in3}} \quad [8-48]$$

In order to achieve as high a gain as possible, we need to make the output impedance of the current source and the input impedance of the voltage source very large compared to r_{o1} . For the current source, let's set $R_{E2} = 1 \text{ k}\Omega$ and $V_{BIAS2} = 9.3 \text{ volts}$, which will give us approximately what we want in terms of bias level. Using these values, we find the incremental parameters of the current source transistor as follows:

$$\begin{aligned} g_{m2} &= 0.077 \text{ A/V} \\ r_{\pi2} &= \frac{h_{fe2}}{g_{m2}} = \frac{175}{0.077} = 2.3 \text{ k}\Omega \\ r_{o2} &= \frac{1}{\eta_{pnp} g_{m2}} = \frac{1}{(1.8 \times 10^{-4})(0.077)} = 72.1 \text{ k}\Omega \\ r_{\mu2} &= h_{fe2} r_{o2} = (175)(72.1 \text{ k}\Omega) = 12.6 \text{ M}\Omega \end{aligned} \quad [8-49]$$

Next, we find the output resistance of the current source:

$$r_{out2} \approx r_{\mu2} \left\| \frac{1 + g_{m2} [r_{\pi2} \parallel R_{E2}]}{g_{o2}} \right\| \approx 3 \text{ M}\Omega \quad [8-50]$$

Note that this value of output resistance is so high that it won't load down the high-gain node significantly.

Next, for the emitter-follower Q_3 , let's assume that we have biased the transistor at a collector current of 1 milliamp. The value of R_{E3} is a little difficult to determine, since we don't know exactly the bias point set at the collector of Q_1 . Let's assume that the collector of Q_1 is at approximately 6 volts (or half the supply voltage). This means that $R_{E3} = 5.3 \text{ k}\Omega$ to set a collector current of 1 milliamp.

We find the incremental parameters of the emitter-follower Q_3 as follows:

$$\begin{aligned} g_{m3} &= \frac{I_{C3}}{\left(\frac{kT}{q} \right)} = 0.038 \text{ A/V} \\ r_{\pi3} &= \frac{h_{fe3}}{g_{m3}} = \frac{200}{0.038} = 5.3 \text{ k}\Omega \\ r_{o3} &= \frac{1}{\eta_{npn} g_{m3}} = \frac{1}{(6.7 \times 10^{-4})(0.038)} = 39.3 \text{ k}\Omega \\ r_{\mu3} &= h_{fe3} r_{o3} = (200)(39.3 \text{ k}\Omega) = 7.9 \text{ M}\Omega \end{aligned} \quad [8-51]$$

Next, we find the input resistance of the emitter-follower:

$$r_{in3} = r_{\mu3} \left\| \left(r_{\pi3} + (1 + h_{fe3}) [R_{E3} \parallel r_{o3}] \right) \approx (7.9 \text{ M}\Omega) \parallel ((200) [(5.3 \text{ k}) \parallel 39.3 \text{ k}]) \approx 835 \text{ k}\Omega \quad [8-52]$$

In order to continue with the analysis, we'll tabulate the small-signal parameters for all transistors in **Table 8-1**.

Table 8-1: Small signal parameters for high-gain amplifier example.

	Q₁	Q₂	Q₃
h_{fe}	200	175	200
g_m	0.077	0.077	0.038
r_{π}	2.6 k Ω	2.3 k Ω	5.3 k Ω
r_o	19.4 k Ω	72.1 k Ω	39.3 k Ω
r_{μ}	3.9 M Ω	12.6 M Ω	7.9 M Ω
C_{π}	39 pF	39 pF	20 pF
C_{μ}	2 pF	2 pF	2 pF

Note that the output resistance r_{o1} of Q_1 is very small compared to the output resistance of the Q_2 current source (r_{out2}) and the input resistance of emitter-follower Q_3 (r_{in3}). Therefore, we expect the gain of this amplifier to be approximately $-1/\eta_{npn}$ or approximately -1400 .

Next, let's estimate the bandwidth using open-circuit time constants. For $C_{\pi1}$, we find the incremental circuit of **Figure 8-25a**. In the following open-circuit time constants calculations we will ignore the r_{μ} of all transistors and include the effects of transistor output resistance r_o only where needed. We note that the open-circuit resistance for $C_{\pi1}$ is zero, because incrementally both sides of $C_{\pi1}$ are grounded.

Next, we note that incrementally $C_{\mu1}$, $C_{\mu2}$, and $C_{\mu3}$ are all in parallel since all three transistors share the high gain node v_{c1} . The open-circuit time constants circuit for this case is shown in **Figure 8-25b**. The incremental resistance to ground at this node is approximately r_{o1} , which is the incremental output resistance of transistor Q_1 . Therefore, this open-circuit time constant is:

$$\tau_{oc1} = r_{o1} (C_{\mu1} + C_{\mu2} + C_{\mu3}) = (19.4 \text{ k}\Omega)(6 \text{ pF}) = 116.4 \times 10^{-9} \text{ sec.} \quad [8-53]$$

The open-circuit resistances for $C_{\pi2}$ and $C_{\pi3}$ are found using the circuits of **Figure 8-25c** and **Figure 8-25d**, respectively. In each case, the open-circuit resistance across each C_{π} is approximately $1/g_m$, since the emitter resistors are relatively large.¹² This results in two more time constants that are small compared to the time constant at the high gain node:

$$\begin{aligned} \tau_{oc2} &\approx \frac{C_{\pi2}}{g_{m2}} \approx \frac{39 \text{ pF}}{0.077} \approx 0.5 \times 10^{-9} \text{ sec.} \\ \tau_{oc3} &\approx \frac{C_{\pi3}}{g_{m3}} \approx \frac{20 \text{ pF}}{0.038} \approx 0.5 \times 10^{-9} \text{ sec.} \end{aligned} \quad [8-54]$$

¹² For details of this calculation, see Chapter 7.

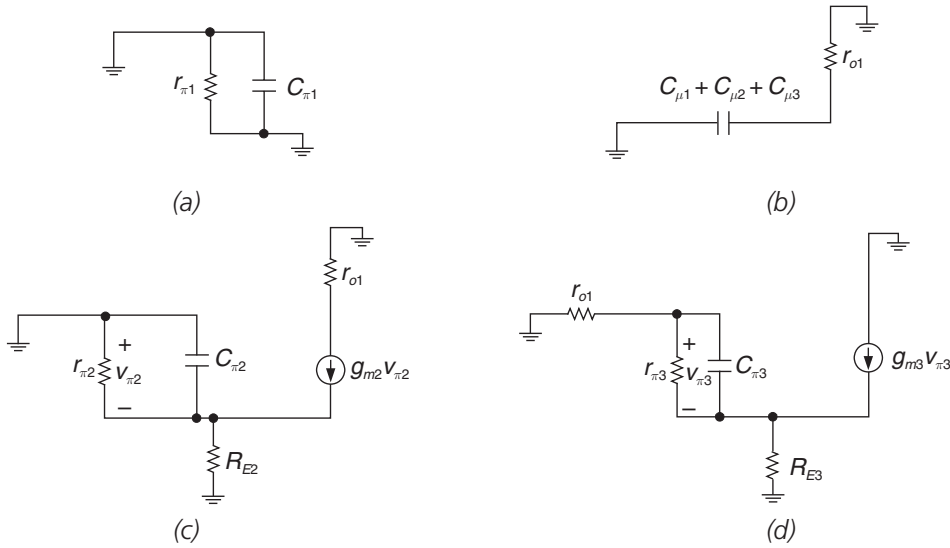


Figure 8-25: Open-circuit time constants circuits for: (a) OCTC circuit $C_{\pi 1}$. (b) OCTC circuit for $C_{\mu 1}$, $C_{\mu 2}$, and $C_{\mu 3}$. (c) OCTC circuit for $C_{\pi 2}$. (d) OCTC circuit for $C_{\pi 3}$.

The open-circuit time constant due to the load capacitor is the incremental output resistance of the Q_3 emitter-follower multiplied by the load capacitor value. This calculation is as follows:

$$r_{out,Q3} \approx R_{E3} \left\| \left(\frac{r_{\pi 3} + r_{o1}}{1 + h_{fe3}} \right) \right\| \approx 120\Omega \quad [8-55]$$

$$\tau_{oc4} = r_{out,Q3} C_L = (120\Omega)(100 \text{ pF}) = 12 \times 10^{-9} \text{ sec.}$$

The sum of the open-circuit time constants for this amplifier is 129.4 nanoseconds, resulting in a bandwidth estimate of:

$$\omega_h \approx \frac{1}{\sum \tau_{oc}} \approx \frac{1}{129.4 \times 10^{-9}} \approx 7.73 \text{ Mrad/sec} \quad [8-56]$$

$$f_h \approx \frac{\omega_h}{2\pi} \approx 1.23 \text{ MHz}$$

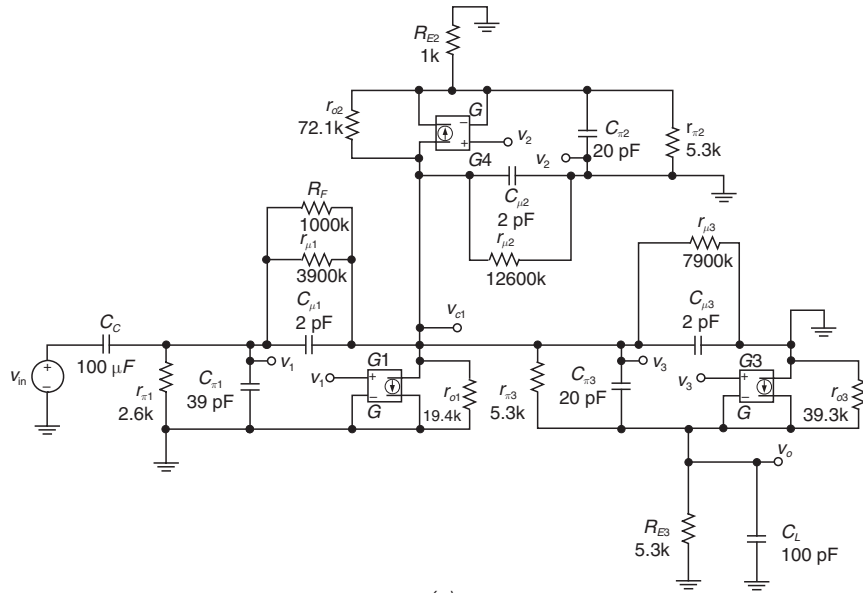
We note that the bandwidth bottleneck is the time constant associated with the capacitance at high-gain node.

In order to estimate the low-frequency breakpoint of this amplifier due to the coupling capacitor, we note that the incremental resistance seen by the coupling capacitor C_C is simply $r_{\pi 1}$. Therefore, the low-frequency breakpoint is:

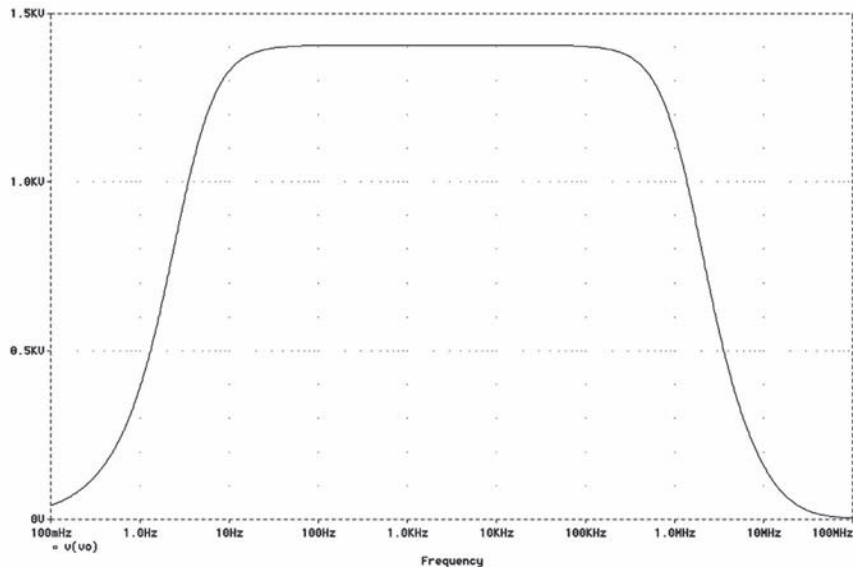
$$\omega_L = \frac{1}{r_{\pi 1} C_C} = \frac{1}{(2600)(100 \times 10^{-6})} = 3.8 \text{ rad/sec} \quad [8-57]$$

$$f_L = \frac{\omega_L}{2\pi} = 0.6 \text{ Hz}$$

A PSPICE analysis (**Figure 8-26**) shows that our estimates of gain and bandwidth are pretty good. Furthermore, we see that the assumption that the emitter-follower has unity gain is justified in this case.



(a)



(b)

Figure 8-26: PSPICE analysis of high-gain amplifier.
(a) Circuit. (b) PSPICE simulation showing frequency response.

Example 8.7: Another high-gain amplifier example (revisited)

In the previous example, we made the assumption that $r_x = 0$ for each transistor. As we'll soon see, this assumption, although it simplifies the math significantly, gives us gain and bandwidth estimates that need to be revisited. Let's redo these estimates, using a value of $r_x = 100\Omega$ for each transistor.

With respect to gain, we expect the finite r_x to reduce the gain, due to its loading effect with $r_{\pi 1}$ and the input resistance of Q_1 . We could march forward and write node equations using the circuit of **Figure 8-27**, but we can approximate the gain reduction with a little forethought. We recognize that the inclusion of r_{x1} means that the voltage $v_{\pi 1}$ is now set by a voltage divider between r_{π} and r_{x1} . Although the feedback resistor of Q_1 is R_F (1000 k Ω) in parallel with $r_{\mu 1}$ (3900 k Ω), the input resistance r_{in} is much lower due to the high negative gain (~ -1400) around this resistance. The input resistance looking into the transistor base, including the effects of R_F' only is approximately:

$$r_{in} \approx \frac{R_F'}{A_V} \approx \frac{796\text{k}}{1400} \approx 570\Omega \quad [8-58]$$

Therefore, the voltage divider that reduces the gain is now:

$$\frac{r_{\pi 1} \parallel r_{in}}{r_{\pi 1} \parallel r_{in} + r_x} = \frac{467}{467 + 100} = 0.82 \quad [8-59]$$

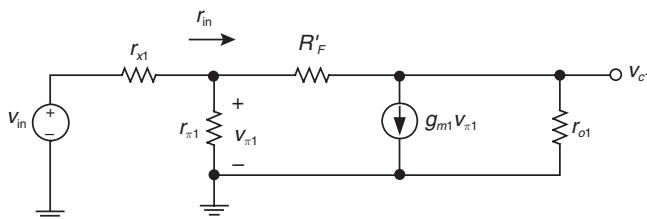


Figure 8-27: Circuit for calculating gain reduction in high-gain amplifier due to finite r_x .

Hence, we expect the gain to be about 82% of that in the previous example,¹³ or approximately -1150 . Therefore, the gain reduction due to a finite r_x is significant.

We will now delve into some detail of the open-circuit time constants calculations, with circuits for each C_{π} and C_{μ} shown in **Figure 8-28**. For $C_{\pi 1}$, we use the circuit of **Figure 8-28a**, where we find the open-circuit resistance to be 96Ω (or approximately the value of r_{x1}) since $r_{x1} \ll r_{\pi 1}$. The open-circuit time constant for $C_{\pi 1}$ is now 3.8 nanoseconds.

¹³ We should iterate the r_{in} calculation with the new value of gain, but you get the idea.

For $C_{\mu 1}$, we use the circuit of **Figure 8-28b**. We note that r_{x1} provides a resistance in the base of transistor Q_1 . We may now expect more Miller multiplication of the capacitance $C_{\mu 1}$. Calculations show that the open-circuit resistance facing $C_{\mu 1}$ is 163.3 k Ω and the resultant open-circuit time constant is 327 nanoseconds. Note that this open-circuit time constant is greater than the total sum of open-circuit time constants for the circuit with $r_x = 0$. Hence, we already know that there is significant bandwidth degradation due to r_x .

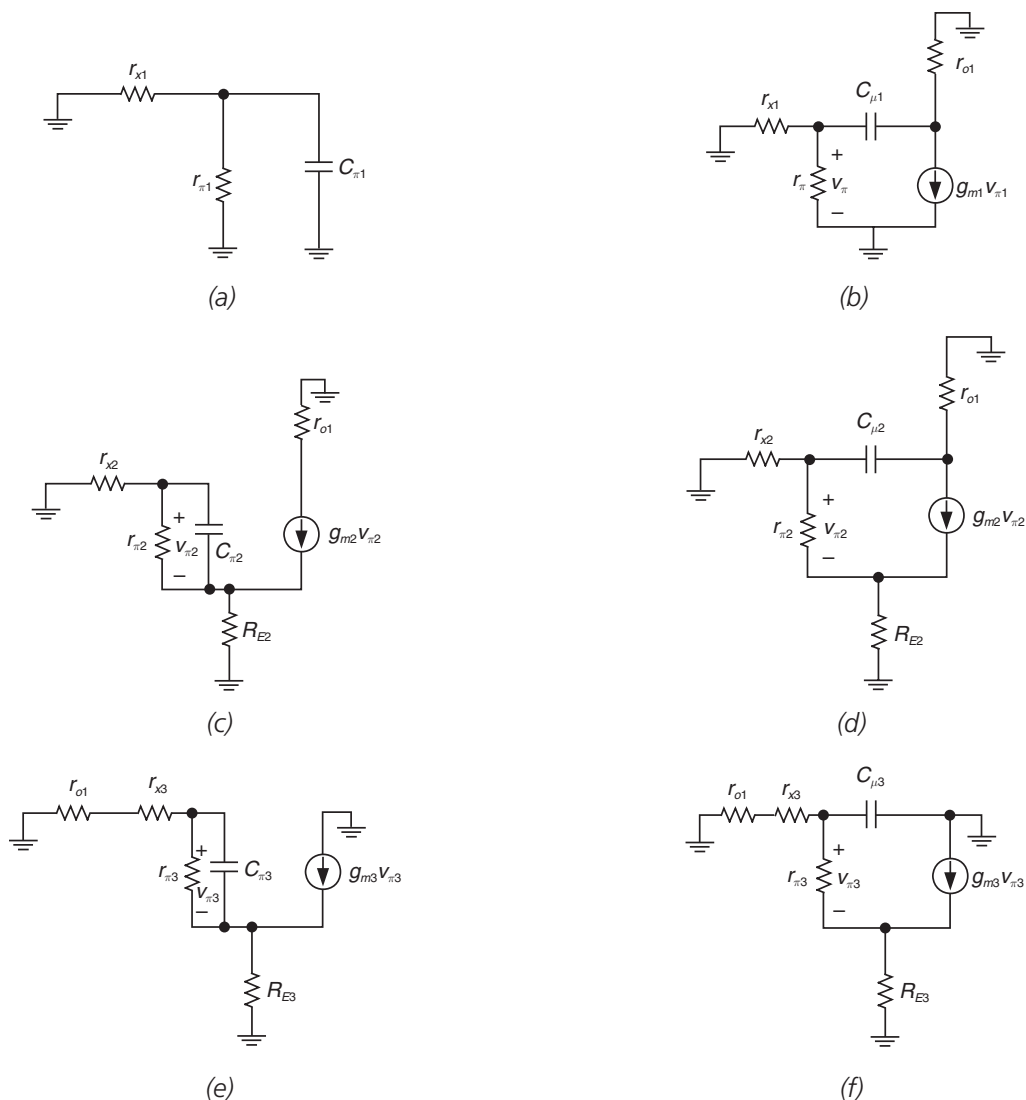


Figure 8-28: Open-circuit time constants circuits for revisited circuit with finite r_x .
 (a) Circuit for $C_{\pi 1}$. (b) $C_{\mu 1}$. (c) $C_{\pi 2}$. (d) $C_{\mu 2}$. (e) $C_{\pi 3}$. (f) $C_{\mu 3}$.

We'll spare the reader the details of further OCTC calculations for the remaining four circuits (**Figure 8-28c – Figure 8-28f**) and the load capacitance, and proceed with the summary in **Table 8-2**. We note that inclusion of r_x has greatly increased the open-circuit time constant for $C_{\mu 1}$ of the common-emitter transistor Q_1 . The overall effect of including r_x is that the OCTC estimate bandwidth of the amplifier is ~ 373 kHz, as compared to a bandwidth of greater than 1 MHz with $r_x = 0$. This result is verified with the PSPICE simulation of **Figure 8-29a** and **Figure 8-29b**. We see that the base spreading resistance reduces the midband gain to approximately $-1,150$ and reduces the bandwidth to ~ 500 kHz. The warning here is to ignore r_x at your own peril!

Table 8-2: Summary of OCTC calculations for high-gain amplifier with $r_x = 100\Omega$.

	Q_1	Q_2	Q_3
R_{oc} for C_π	96.3Ω	14Ω	120Ω
R_{oc} for C_μ	163.3 k Ω	21.4 k Ω	19.1 k Ω
τ_{oc} for C_π	3.8 ns	0.5 ns	2.4 ns
τ_{oc} for C_μ	327 ns	43 ns	38 ns
τ_{oc} for C_L	12 ns		
Sum of OCTCs	427 ns		
Estimate of ω_h	2.34 Mrad/sec		
Estimate of f_h	373 kHz		

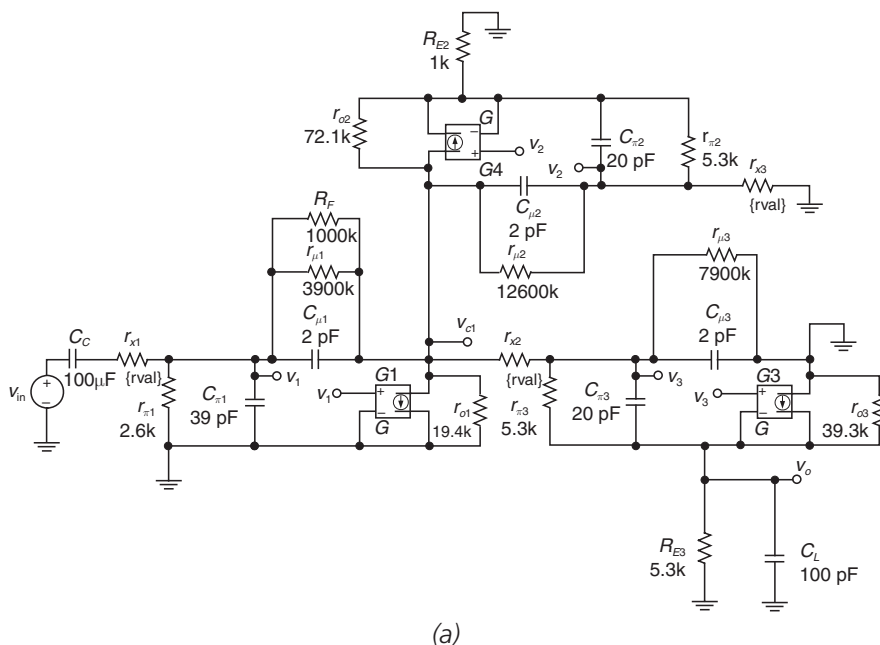
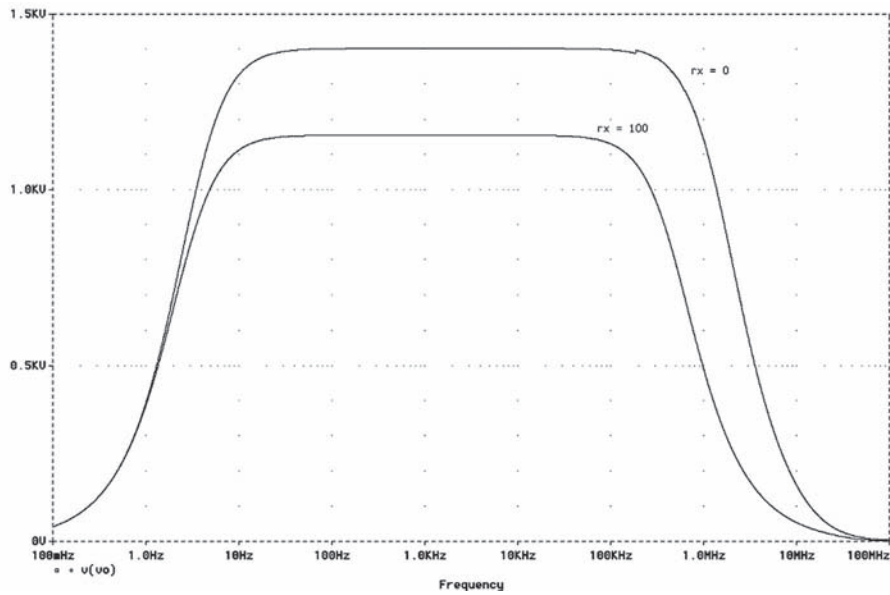


Figure 8-29: PSPICE analysis of high-gain amplifier with $r_x = 100\Omega$.
(a) Circuit (continued on following page).



(b)

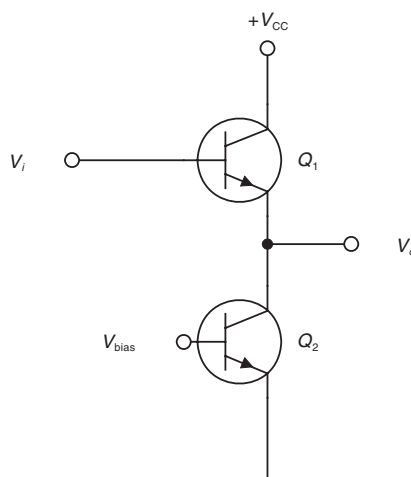
Figure 8-29 (continued): (b) Comparison of frequency responses of amplifiers with $r_x = 0$ and $r_x = 100\Omega$ for each transistor. Note that a finite r_x both decreases the gain and bandwidth compared to $r_x = 0$.

Chapter 8 Problems

Problem 8.1

Using the extended hybrid-pi model, determine the low-frequency input resistance of the emitter-follower connection in **Figure 8-30** as a function of transistor parameters and quiescent operating levels. Assume both transistors are identical and that the collector of Q_1 is tied to $+V_{CC}$.

Figure 8-30: Circuit for Problems 8.1 and 8.2.



Problem 8.2

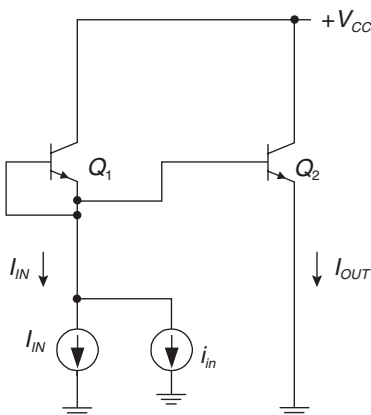
Again using the extended hybrid-pi model and the circuit of **Figure 8-30**, find the output resistance seen at the collector of Q_1 , as a function of transistor parameters and quiescent operating conditions.

Problem 8.3

Using open-circuit time constants, estimate the small-signal speed of response of the current mirror shown in **Figure 8-31**. Assume that each transistor is a 2N3906 biased at 10 milliamps with $V_{CB} = 6$ volts. Assume that $r_x = 0\Omega$ and that incrementally I_{in} is a very high impedance and I_{out} presents a very low impedance.

- Draw the small-signal model.
- Estimate the bandwidth using open-circuit time constants.
- Simulate using PSPICE and compare your PSPICE result to the OCTC estimate.

Figure 8-31: Circuit for Problems 8.3 and 8.4.



Problem 8.4

Let's revisit the current mirror circuit of Problem 8.3, but this time assume that the transistors in question instead of having zero base resistance have base resistance in the range $r_x = 100\Omega$ to 200Ω . Unfortunately, with a finite value of r_x , the open-circuit time constants problem is very difficult in this case. Resimulate using PSPICE and find the bandwidth with this range of finite value of r_x . Comment on how this solution compares to the solution with $r_x = 0$.

Problem 8.5

Design a transistor amplifier with gain magnitude of $|A_v| = 10000$, using high-gain techniques. You're driving a capacitive load with your amplifier, and assume that you need an emitter-follower at the output to isolate the high gain node from the capacitive load. Assume that you have at your disposal transistors with $\eta_{npn} = 6.7 \times 10^{-4}$ and $\eta_{pnp} = 1.8 \times 10^{-4}$. Don't worry about biasing details (you can assume that your collector currents will magically run at the correct bias levels. Ignore r_x in all your calculations and assume that $h_{fe,npn} = 200$ and $h_{fe,pnp} = 175$. Simulate the gain of your circuit using SPICE. Assuming that $C_\pi = 0$ and $C_\mu = 2$ pF for each transistor, estimate the bandwidth of your circuit using OCTCs and confirm using SPICE.

References

- Early, J. M., "Effects of Space-Charge Layer Widening in Junction Transistors," *Proceedings of the I.R.E.*, 1952, pp. 1401–1406.
- Gray, P. E., DeWitt, D., Boothroyd, A. R., and Gibbons, J. F., *Physical Electronics and Circuit Models of Transistors*, Semiconductor Electronics Education Committee, volume 2, John Wiley, 1964.
- Gray, Paul R., Hurst, Paul J., Lewis, Stephen H., and Meyer, Robert G., *Analysis and Design of Analog Integrated Circuits*, 4th edition, John Wiley, 2001.
- Grebene, Alan B., *Bipolar and MOS Analog Integrated Circuit Design*, John Wiley, 1984.
- Hart, B. L., "Modeling the Early Effect in Bipolar Transistors," *IEEE Journal of Solid State Circuits*, vol. 18, no. 1, February 1983, pp. 139–140.
- Huiting, Chen, Whiteside, F., and Geiger, R., "Current Mirror Circuit with Accurate Mirror Gain for Low β Transistors," *Proceedings of the 2001 IEEE International Symposium on Circuits and Systems*, (ISCAS 2001), May 6–9, 2001, pp. 536–539.
- Joardar, K., "A New Approach for Extracting Base Width Modulation Parameters in Bipolar Transistors," *Proceedings of the 1994 Bipolar/BiCMOS Circuits and Technology Meeting*, October 1994, pp. 140–143.
- Kimura, K., "Low Voltage Techniques for Bias Circuits," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 5, May 1997, pp. 459–465.
- Liou, J. J., "Comments on 'Early Voltage in Very-Narrow-base Bipolar Transistors,' by D. J. Roulston" *IEEE Electron Device Letters*, vol. 11, no. 5, May 1990, p. 236.
- Mahattanakul, J., Pookaiyaudom, S., and Toumazou, C., "Understanding Wilson Current Mirror via the Negative Feedback Approach," *Proceedings of the 2001 IEEE International Symposium on Circuits and Systems*, (ISCAS 2001), May 6–9, 2001, pp. 532–535.
- McAndrew, C. C., and Nagel, L. W., "SPICE Early Modeling [Bipolar Transistors]," *Proceedings of the 1994 Bipolar/BiCMOS Circuits and Technology Meeting*, October 10–11, 1994, pp. 144–147.
- Roberge, James K., *Operational Amplifiers: Theory and Practice*, John Wiley, 1975.
- Roulston, D. J., "Early Voltage in Very-Narrow-base Bipolar Transistors," *IEEE Journal of Solid State Circuits*, vol. 11, no. 2, February 1990, pp. 88–89.
- Rucker, L. M., "Monolithic Bipolar Diodes and Their Models," *IEEE Circuits and Devices Magazine*, vol. 7, no. 2, March 1991, pp. 26–31.
- Soclof, Sidney, *Analog Integrated Circuits*, Prentice-Hall, 1985.
- Thompson, Marc T., "Tips for Designing High-Gain Amplifiers," *Electronic Design*, May 16, 1994, pp. 83–90.
- van Kessel, J., and van de Plassche, R. J., "Integrated Linear Basic Circuits," *Philips Technical Review*, vol. 32, no. 1, 1971, pp. 1–10.
- Yuan, J. S., and Liou, J. J., "An Improved Early Voltage Model for Advanced Bipolar Transistors," *IEEE Transactions on Electron Devices*, vol. 38, no. 1, January 1991, pp. 179–182.

Introduction to MOSFET Devices and Basic MOS Amplifiers

In This Chapter

- We next take a detour from the world of bipolar transistors and enter the world of metal-oxide semiconductor field-effect (MOSFET¹) transistors. The basic signal MOS gate is discussed, followed by a discussion of MOS amplifiers. The incremental model of the MOS transistor is shown, and it is used in a design example where gain and bandwidth are calculated for a MOS amplifier.

Some Early History of Field-Effect Transistors

The invention of the metal-oxide semiconductor field-effect transistor (MOSFET) pre-dates the bipolar transistor. An excerpt from one U.S. patent granted in 1933 to Dr. Julius Lilienfeld is shown in **Figure 9-1**. In three patents, Dr. Lilienfeld gave structures of the MOSFET, MESFET and other MOS devices, but he wasn't able to build any working FETs, underscoring the difficulty in fabricating practical semiconductor devices at that time. In fact, it was not until the 1960s that the first commercially successful FET devices were manufactured.

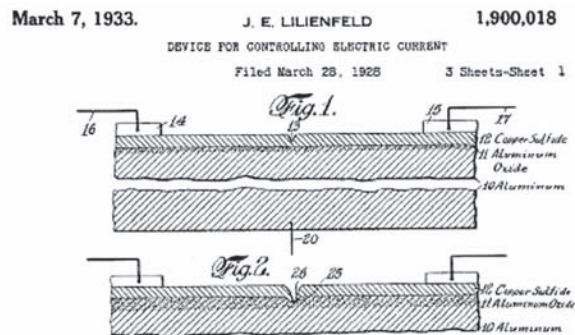


Figure 9-1: Excerpt from Lilienfeld's U.S. patent 1,900,018² (1933) showing two different versions of the MOSFET. In Figure 1 from the patent, terminal #16 is the source, terminal #15 is the drain and terminal #20 is the gate connection.

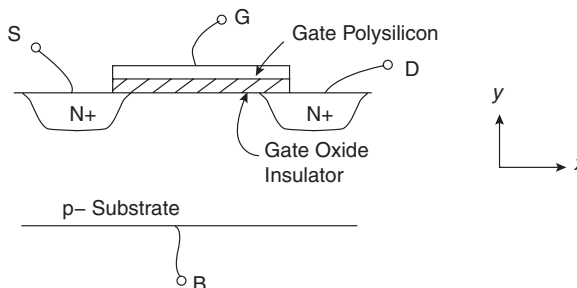
¹ We will not cover the JFET (junction field-effect transistor).

² Lilienfeld had three patents in succession covering basic MOS transistor structures. They are U.S. patent #1,745,175 (filed 10/8/26, granted 1/18/30); U.S. patent #1,877,140 (filed 12/8/28, granted 9/13/32) and U.S. patent #1,900,018 (filed 3/28/28, granted 3/7/33).

Qualitative Discussion of Basic MOS Devices

A basic N-channel lateral MOS gate is shown in **Figure 9-2**. This is a lateral device because current flow is in the $-x$ direction laterally across the surface of the device.³ An N-channel MOS device starts with a lightly doped p- substrate. N-type source and drain regions are added at either ends of the channel. Next, an insulating oxide layer is grown on the surface. A gate connection is isolated from the p- substrate by this oxide layer that has a very low electrical conductivity. This oxide layer forms a gate capacitance. A fourth terminal (called bulk or substrate) is connected to the lowest potential in the circuit. The four terminals shown are source (S), gate (G), drain (D) and substrate, or bulk (B).

Figure 9-2: Basic N-channel MOS device. Device has width W into the page. The terminals are source (S), gate (G), drain (D) and the substrate (B, for "bulk").



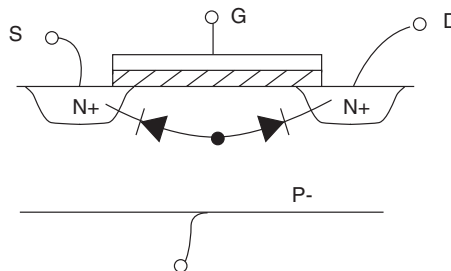
The gate contact forms a capacitance⁴ per unit area through the oxide layer, of value:

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} \quad [9-1]$$

where ϵ_{ox} is the dielectric permittivity of the oxide layer and t_{ox} is the thickness of the oxide layer.

Now, in a thought experiment, let's consider the equivalent circuit of the MOSFET, when it is OFF. If the gate terminal is not connected, the source-channel junction and drain-channel junction path each behave as diodes, and these diodes are back-to-back as shown. No current can flow through this device (**Figure 9-3**), other than diode leakage current.

Figure 9-3: N-channel MOS device, showing back-to-back diodes. With the gate disconnected there is no current flow in the MOS device.

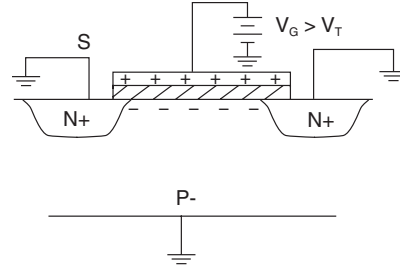


³ Power MOS devices are generally vertical devices; current flow is vertically through the device. Signal MOS devices have current flow laterally across the surface.

⁴ ϵ_{ox} has units of farads/meter, so C_{ox} has units of capacitance per square meter.

In a second thought experiment, let's ground the source and the drain, and apply a positive voltage to the gate (**Figure 9-4**). This gate voltage produces an electric field in the oxide layer that originates on positive charges in the gate electrode and terminates on negative charges in the channel. If we apply sufficient positive voltage to the gate, we will attract sufficient mobile electrons to the channel underneath the gate to form an *inversion layer*. Basically, at this magic voltage (called the *threshold voltage* V_T) we have caused the P-material under the gate to behave as if it were n-material, due to this inversion layer that we have formed. Between the N+ islands, we see an inversion layer that behaves as an N-region. Below a gate-source voltage of V_T , the MOSFET is essentially OFF; for $V_{GS} > V_T$ the MOSFET can support current flow.

Figure 9-4: N-channel MOS device with positive gate potential greater than the threshold voltage (V_T) applied, forming an inversion layer.



In order to figure out the details of the shape of the MOS transistor V/I curve, we have to resort to a mathematical model.

Figuring Out the V/I Curve of a MOS Device

Let's consider a more detailed model of the N-channel MOSFET (**Figure 9-5**) in the hopes of generating the ideal MOS V/I curve. If there is current flowing in the MOSFET channel, we would expect a varying voltage along the length of the channel. Let's call this channel voltage $V_c(x)$. Remember that no current will flow for gate to source (V_{GS}) voltages less than the threshold voltage V_T . We can find the charge per unit area under the gate in the channel as:

$$Q_c(x) = -C_{ox}(V_{GS} - V_T - V_c(x)) \quad [9-2]$$

Note that a mobile gate charge exists only if the gate-source voltage exceeds the threshold voltage. The gate charge $Q_c(x)$ is the charge density (Coulombs/m²) at position x along the gate.

The drain current is equal to:⁵

$$I_D = -WQ_c(x)v(x) \quad [9-3]$$

where W is the width of the MOS device into the paper, and $v(x)$ is the velocity of the charges through the channel.

⁵ Note that the units work out. Current has units of coulombs/second; W has units of meters, $Q_c(x)$ has units of coulombs/m² and velocity has units of meter/second.

If we assume a low electric field (and hence a long length device), the drift velocity is linearly related to the electric field as:

$$v(x) = \mu_n E_x = \mu_n \frac{dV_c(x)}{dx} \quad [9-4]$$

We can then express the drain current as:

$$I_D = WC_{ox} [V_{GS} - V_c(x) - V_T] \mu_n \frac{dV_c(x)}{dx} \quad [9-5]$$

We can bring the derivative dx to the left-hand side, resulting in:

$$I_D dx = W \mu_n C_{ox} [V_{GS} - V_c(x) - V_T] dV_c(x) \quad [9-6]$$

Next, let's integrate both sides. We integrate x over the length of the channel from 0 to L , and we integrate the channel voltage $V_c(x)$ from 0 to V_{GS} .

$$\int_0^L I_D dx = \int_0^{V_{GS}} W \mu_n C_{ox} [V_{GS} - V_c(x) - V_T] dV_c(x) \quad [9-7]$$

Following out this integral results in:

$$\begin{aligned} I_D L &= \mu_n W C_{ox} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \\ &\Downarrow \\ I_D &= \mu_n C_{ox} \left(\frac{W}{L} \right) \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \end{aligned} \quad [9-8]$$

This is the drain current equation for the ideal MOSFET, valid for gate-source voltage greater than the threshold voltage, or $V_{GS} > V_T$.

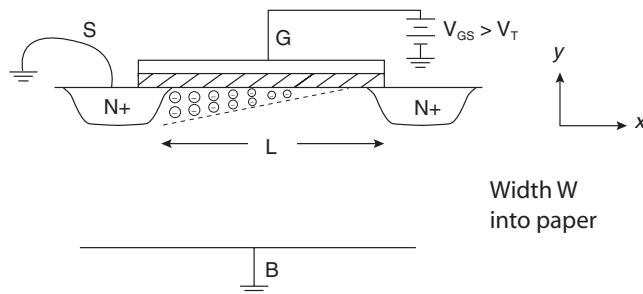
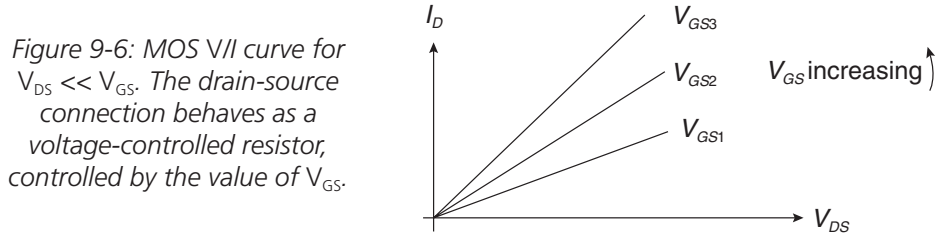


Figure 9-5: N-channel MOS device with positive gate potential ($>V_T$) applied, showing an inversion layer under the gate.

For very small V_{DS} ($V_{DS} \ll V_{GS}$) and with $V_{GS} > V_T$, this device behaves like a voltage-controlled resistor:

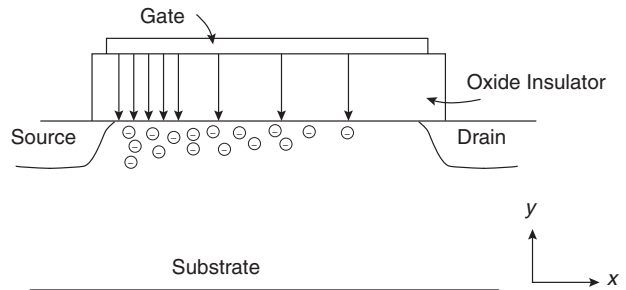
$$I_D \approx \mu_n C_{ox} \left(\frac{W}{L} \right) ((V_{GS} - V_T) V_{DS}) \quad V_{DS} \ll V_{GS} \quad [9-9]$$

As V_{DS} increases, you get higher current because there is a higher lateral electric field sweeping carriers through the channel. Current also increases with higher $(V_{GS} - V_T)$ since this increases the charge under the gate available to support current. This region of operation is shown in **Figure 9-6**.



If V_{DS} is equal to or higher than $V_{GS} - V_T$, there is sufficient drain-source voltage to “pinch off” the channel at the drain end of the channel. This means that the gate charge at the end of the channel approaches zero as shown in **Figure 9-7**, and we achieve maximum drain current supported by the available V_{GS} . If V_{DS} increases, to first order there is no further increase in drain current. In the MOS device, this is the so-called saturation⁶ region.

Figure 9-7: Qualitative view of gate charge and electric field inside oxide layer for MOS device operating in the “pinchoff” region with $V_{DS} > V_{GS} - V_T$. At the drain end of the channel there is no mobile gate charge and the drain is pinched off. (Dimensions distorted for clarity).



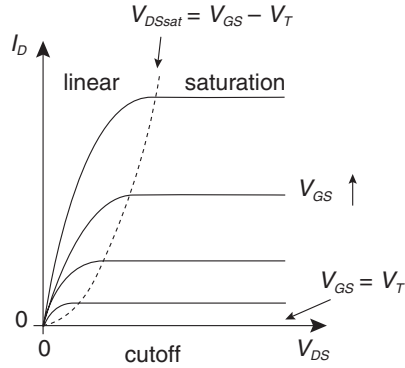
For $V_{DS} > (V_{GS} - V_T)$ we find that the drain current saturates. Indeed, we can find the drain current in the saturation region by setting $V_{DS} = (V_{GS} - V_T)$ in the drain current equation. This results in the drain current in saturation:

$$I_D = \frac{\mu_n C_{ox}}{2} \frac{W}{L} (V_{GS} - V_T)^2 \quad [9-10]$$

The ideal MOS V/I curves are shown in **Figure 9-8**.

⁶ Don't get this confused with the bipolar transistor, where the saturation region occurs for V_{CE} less than $V_{CE, sat}$.

Figure 9-8: MOS V/I curve showing linear, saturation and cutoff regions of operation.



MOS Small-Signal Model (Low Frequency)

A MOS device is useful as an amplifier only if operated in the saturation region, or with $V_{DS} > (V_{GS} - V_T)$ and with $V_{GS} > V_T$. In this region of operation, amplification can occur since the drain current varies with V_{GS} but (to first-order) this drain current does not vary with V_{DS} . This region of operation is analogous to the forward-active region in bipolar transistors, where amplification can occur.

If we are operating in the MOS saturation region, in order to find small-signal variation of drain current with gate-source voltage, we need to solve for transconductance:

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_T) = \sqrt{2\mu_n C_{ox} \frac{W}{L} I_D} \quad [9-11]$$

This expression means that the transconductance of a MOS device scales as the square-root of bias current I_D . Compare this expression to that of the bipolar transistor, where transconductance scales as the first power of collector bias current.

The MOS transistor also has a finite output impedance due to *channel-length modulation*, which is analogous to base-width modulation in the bipolar transistor. The width of the drain depletion region varies as V_{DS} varies, with the effect that drain current also varies with V_{DS} . Analogous to our development of the bipolar transistor, we can modify the large-signal V-I characteristic of the MOSFET by including a parameter λ that helps model the finite output resistance of the device:

$$I_D = \frac{\mu_n C_{ox} W}{2 L} (V_{GS} - V_T)^2 (1 + \lambda V_{DS}) \quad [9-12]$$

Incrementally, we can model this effect as an output resistance r_o across the drain-source terminals. The output resistance in the small-signal model can be found as:

$$r_o = \frac{1}{\frac{\partial I_D}{\partial V_{DS}}} = \lambda I_D \quad [9-13]$$

The source-bulk terminals act as a second set of input terminals. In some cases, we will tie the source and bulk terminals together, and therefore we can ignore the effects of this second input. However, if the bulk is *not* tied to the same potential as the source, we need to consider this effect. The result is to include in our small-signal model a transconductance to model this *back-gate effect*, or:

$$g_{mb} = \frac{\partial I_D}{\partial V_{BS}} \quad [9-14]$$

Putting this all together, we find the small-signal model of the MOS transistor at low frequencies, in the saturation region (**Figure 9-9**).

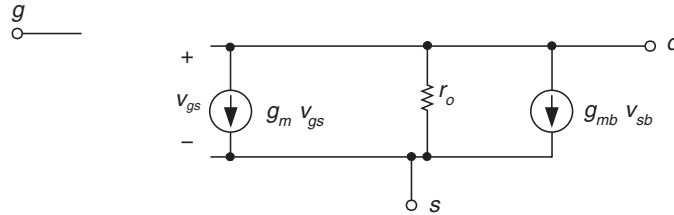


Figure 9-9: MOS low-frequency small-signal model in the saturation region, showing gate (g), drain (d), and source (s) terminals. The bulk (b), or substrate connection is also assumed, but not shown. The $g_{mb}v_{sb}$ generator source causes a current flow proportional to source-bulk voltage v_{sb} .

MOS Small-Signal Model (High Frequency)

The MOS small-signal model at high frequencies has capacitive elements modeling gate-source capacitance (C_{gs}), gate-drain (C_{gd}), source-substrate (C_{sb}) and drain-substrate (C_{db}). Capacitance C_{gs} models the effect of the charge under the gate. C_{gd} models the effect of the gate oxide overlap over the drain region. C_{db} and C_{sb} are depletion capacitances between drain-substrate and source-substrate, respectively. A MOS transistor high-frequency model small-signal model is shown in **Figure 9-10**.

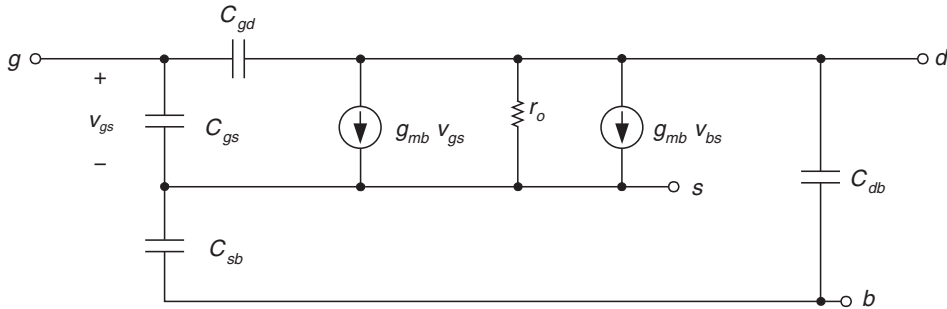


Figure 9-10: MOS high-frequency small-signal model showing gate-source capacitance (C_{gs}), gate-drain capacitance (C_{gd}), source-bulk capacitance (C_{sb}) and drain-bulk capacitance (C_{db}).

Basic MOS Amplifiers

The topologies of MOS amplifiers are similar to that of bipolar transistors, as shown in **Table 9-1**.

Table 9-1: Comparison of BJT and MOS amplifiers.

BJT	MOS
Emitter follower	Source follower
Common emitter	Common source
Common base	Common gate
BJT differential amplifier	MOS differential amplifier

When analyzing MOS amplifiers, there are several differences to consider, as compared to bipolar amplifier analysis:

- The incremental input resistance of a MOS device is very high. In the BJT world, the input resistance at the base is limited by r_π and $(1 + \beta_o)R_E$.
- A MOSFET is a 4-terminal device. In the MOS world, the connection to the substrate (or bulk) must also be considered. Furthermore, there are extra parasitic elements to the substrate (C_{sb} and C_{db}) that must be considered.
- The MOS device has another dependent current generator that must be considered, due to the back-gate effect. In effect, the gate-bulk voltage acts as an extra gate.

Source follower

A source-follower buffer is shown in **Figure 9-11a**. This circuit has a high input impedance, a low output impedance, and a gain close to 1.0. The small-signal low-frequency model⁷ is shown in **Figure 9-11b**. Using this model we can find the small-signal gain of this buffer as by doing KCL at the output node, we find:

$$g_m v_{gs} - v_o g_o = 0 \rightarrow g_m (v_i - v_o) - v_o g_o = 0 \quad [9-15]$$

Simplifying, we find:

$$\frac{v_o}{v_i} = \frac{g_m}{g_m + g_o} \quad [9-16]$$

This gain is very close to +1 if the transconductance of the MOS device is large compared to $1/r_o$, or if $g_m \gg g_o$.

⁷ In this example we ignore the back-gate effect, so $g_{mb} = 0$. We also note that $v_{gs} = v_i - v_o$.

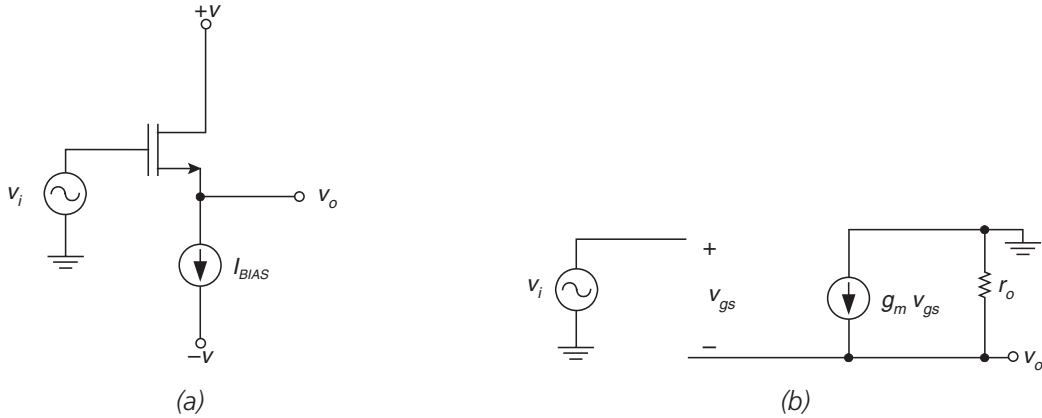


Figure 9-11: Source follower. (a) Circuit. (b) Small-signal low-frequency model.

Common-source amplifier

A common-source amplifier is shown in **Figure 9-12a**. This circuit has a high input impedance and provides voltage gain with a voltage inversion. The small-signal low-frequency model⁸ is shown in **Figure 9-12b**. Using this model we find the small-signal gain:

$$\frac{v_o}{v_i} = -g_m (R_L \parallel r_o) \quad [9-17]$$

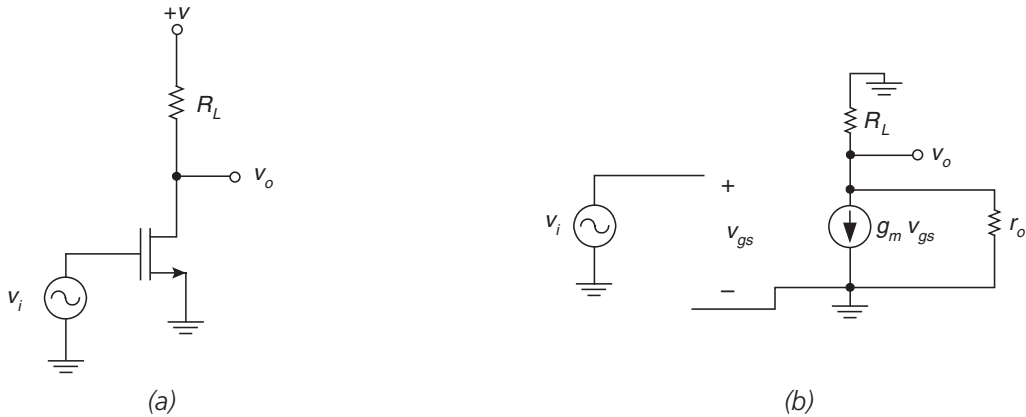


Figure 9-12: Common-source amplifier with resistive load.
(a) Circuit, omitting biasing details for simplicity. (b) Small-signal low-frequency model

⁸ Again, ignoring the back-gate effect.

If we load the common-source amplifier with an ideal current source instead, we have the circuit of **Figure 9-13**. The gain of this amplifier is:

$$\frac{v_o}{v_i} = -g_m r_o \quad [9-18]$$

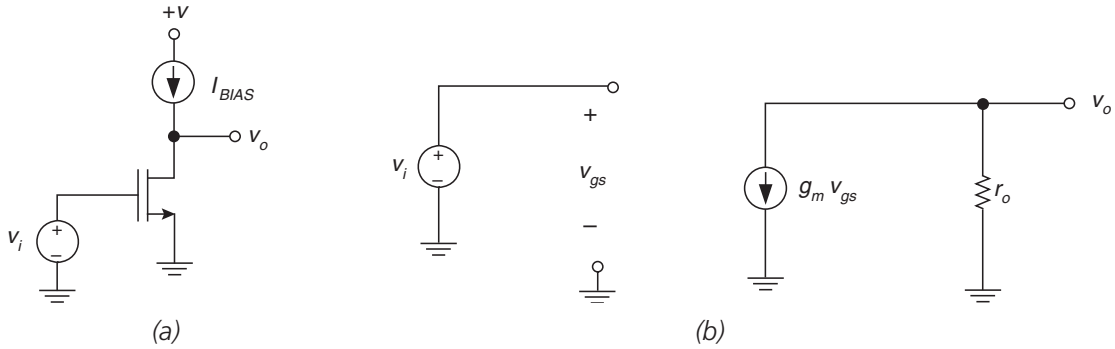


Figure 9-13: Common-source amplifier with current source load.
 (a) Circuit, omitting biasing details for simplicity and assuming an ideal current source with infinite incremental output resistance. (b) Small-signal low-frequency model assuming negligible back-gate effect (i.e., assuming $g_{mb} = 0$).

Common-gate amplifier

A common-gate amplifier is shown in **Figure 9-14**. We can solve for v_{gs} by using the incremental circuit of **Figure 9-14b** and by summing currents at the source:

$$\frac{v_i + v_{gs}}{R_i} + g_m v_{gs} = 0 \Rightarrow v_{gs} = -v_i \left(\frac{1}{1 + g_m R_i} \right) \quad [9-19]$$

The input-output transfer function is found as follows:

$$\begin{aligned} v_o &= -g_m v_{gs} R_L = v_i \left(\frac{g_m R_L}{1 + g_m R_i} \right) \\ \therefore \frac{v_o}{v_i} &= \frac{g_m R_L}{1 + g_m R_i} \end{aligned} \quad [9-20]$$

Note that the gain is approximately R_L/R_i if $g_m R_i \gg 1$.

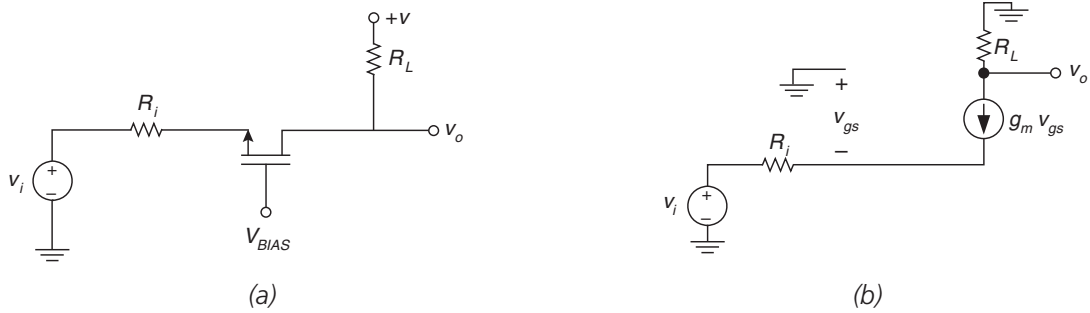


Figure 9-14: Common-gate amplifier with resistive load. (a) Circuit, omitting biasing details for simplicity. (b) Small-signal low-frequency model assuming negligible back-gate effect (i.e., assuming $g_{mb} = 0$).

We can use the common-gate amplifier to augment the common-source amplifier in a cascode configuration (**Figure 9-15a**). In this widely used configuration, common-source transistor M_1 is loaded by common-gate buffer M_2 . The voltage-controlled current generated at the drain of M_1 is buffered by M_2 and flows through the load resistor R_L . The source of M_2 presents a low impedance to the drain of M_1 , hence eliminating the Miller effect.

Another variation on the theme is the *folded cascode* of **Figure 9-15b**. The folded cascode is topologically similar to the basic cascode in that it consists of a common-source transistor feeding a common-gate buffer transistor with current. The P-channel common-gate provides buffering between M_1 and the current source load.

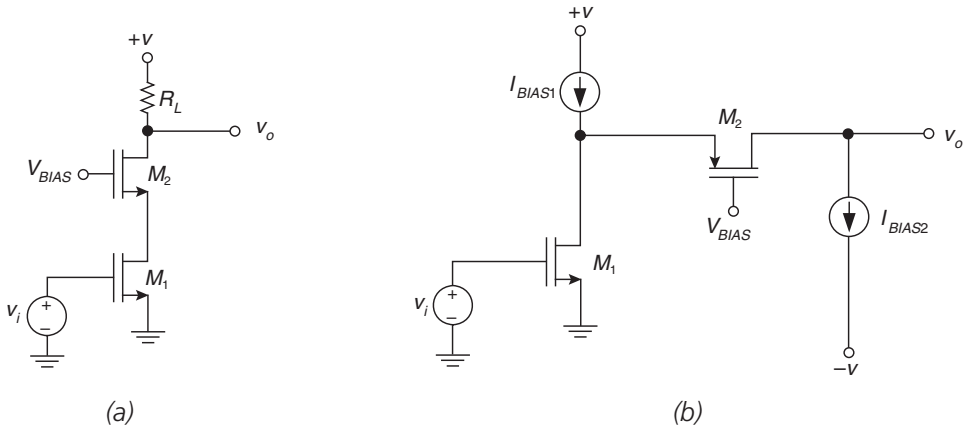


Figure 9-15: MOS cascode amplifiers. (a) MOSFET M_1 is a common-source amplifier stage. MOSFET M_2 is a common-gate connection providing buffering between the common-source amplifier and load resistor R_L . (b) Folded cascode where M_1 is an N-channel MOSFET, and M_2 is a P-channel MOSFET.

MOS current mirrors

A basic CMOS current mirror is shown in **Figure 9-16a**. This building block works on the same basic premise as a bipolar mirror—to first order, CMOS gates with the same V_{GS} will have the same drain current I_D , provided that both devices operate in the saturation region.

In order to estimate the small-signal bandwidth of the MOS mirror, we refer to **Figure 9-16b**. Note that all capacitances appear in parallel to one another. The resistance across the capacitance nodes is $\sim 1/(2g_m)$. Therefore, using open-circuit time constants, an estimate of the small-signal bandwidth of this circuit is:

$$\omega_h \approx \frac{2g_m}{C_{gs1} + C_{db1} + C_{gs2} + C_{gd2}} \quad [9-21]$$

We also note that the low-frequency incremental output resistance of the simple mirror can be quite modest, approximately r_o .

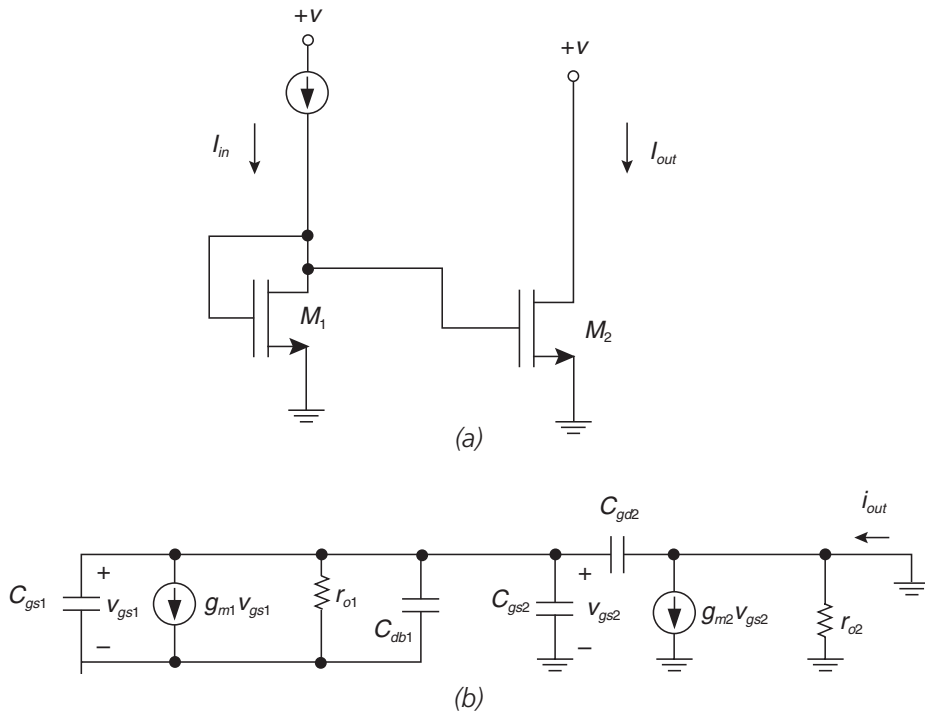


Figure 9-16: Basic MOS current mirror. (a) Circuit. (b) High-frequency incremental model.

The Wilson current mirror (**Figure 9-17a**) uses negative feedback to increase the output resistance of the mirror. We can see how this operates using the low-frequency small-signal model of **Figure 9-17b**. We'll assume that all the MOS devices are identical and biased at the same drain currents, and hence $g_{m1} = g_{m2} = g_{m3}$ and $r_{o1} = r_{o2} = r_{o3}$. The diode-connected MOS device M_2 has been replaced by its output resistance ($1/g_{m2}$). We've added a test current source i_t and

we'll now calculate the test voltage v_t in order to find the incremental output resistance at the M_3 drain node.

Applying KVL around the loop containing the test voltage source results in:

$$v_t = i_x r_o + v_b = i_x r_{o3} + \frac{i_t}{g_{m2}} \quad [9-22]$$

The unknown current i_x is found by:

$$i_x = i_t - g_{m3} v_a \quad [9-23]$$

The unknown voltage v_a is found by:

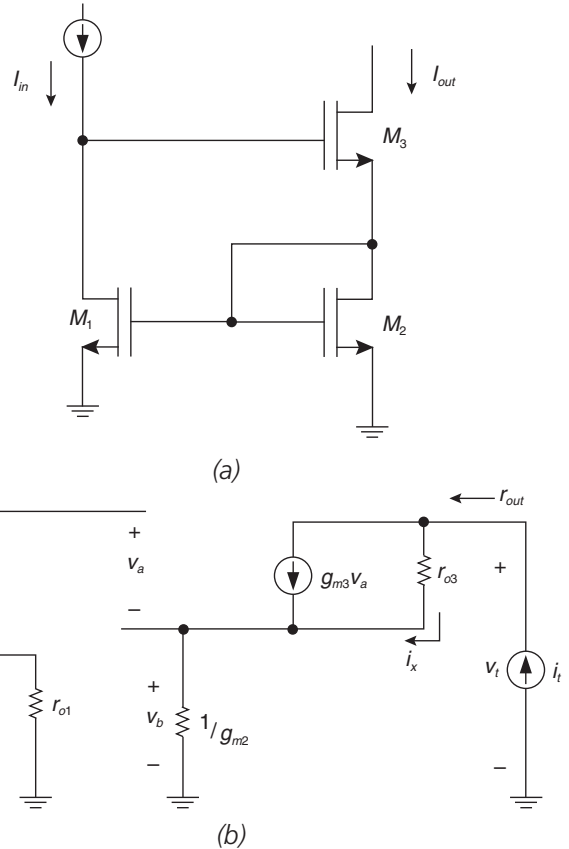
$$v_a = -g_{m1} v_b r_{o1} = -g_{m1} \left(\frac{i_t}{g_{m2}} \right) r_{o1} \quad [9-24]$$

Solving for the output resistance, we find:

$$r_{out} = \frac{v_t}{i_t} \approx \left(1 + \frac{g_{m1} g_{m3}}{g_{m2}} r_{o1} \right) r_{o3} \quad [9-25]$$

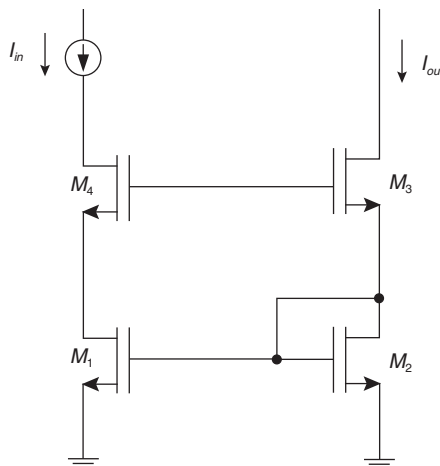
Note that this output resistance is significantly higher than that of the simple current mirror.

Figure 9-17. Wilson current mirror.
(a) Circuit. (b) Low-frequency incremental model.



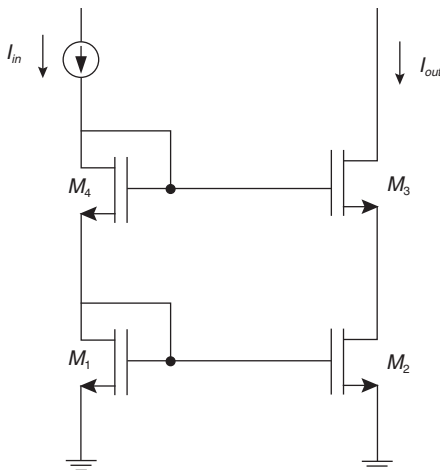
Another variation on the Wilson mirror is the “improved” Wilson mirror (**Figure 9-18**). MOS device M_4 is added, which ensures that the drain-source voltages of M_1 and M_2 are equal. This reduces gain errors due to the finite output resistance of M_1 and M_2 .

Figure 9-18: “Improved” Wilson current mirror.



Yet another MOS mirror that provides a high output impedance is the cascode current mirror (**Figure 9-19**). Note that output transistor M_3 drives into the drain terminal of M_2 . This insures that the output resistance seen at the drain of M_3 is high.

Figure 9-19: Cascode current mirror.



Example 9.1: MOS amplifier design example

Let's work out a MOS design example illustrating the use of MOS models and open-circuit time constants for bandwidth analysis. Let's assume that we need a gain magnitude of 10 (i.e., 20dB) and a -3dB bandwidth in excess of 350 MHz. The input source has a series resistance of 1 k Ω and that the amplifier drives a 1-picofarad load capacitance. Ignore the back gate effect (i.e., assume that $g_{mb} = 0$). Assume that the bulk (i.e., substrate) connections of the transistors are tied to ground. The MOS transistors have the following parameters:

- $g_m = 0.01$ A/V
- $r_o = 3$ k Ω
- Gate-source capacitance $C_{gs} = 0.2$ pF
- Gate-drain capacitance $C_{gd} = 0.05$ pF
- Source-substrate capacitance $C_{sb} = 0.02$ pF
- Drain-substrate capacitance $C_{db} = 0.02$ pF

TRY #1: Common-source amplifier

A first-cut design for this amplifier is shown in **Figure 9-20a**, with low-frequency small-signal model shown in **Figure 9-20b**. We can use the low-frequency small-signal model to find the minimum value of the load resistor R_L in order to meet the gain specification, as follows:

$$\frac{v_o}{v_i} = -g_m R_L \parallel r_o \Rightarrow R_L > 1500\Omega \quad [9-26]$$

We'll use $R_L = 1.6$ k Ω so that there's a little bit of extra gain, and also to account for the fact that we've ignored the back-gate effect.

We'll find the bandwidth using the high-frequency small-signal model of **Figure 9-20c**. Note that the load (C_L) and drain-substrate (C_{db}) capacitances are in parallel. The source-substrate (C_{sb}) capacitance is not included since the source and substrate are both incrementally grounded.

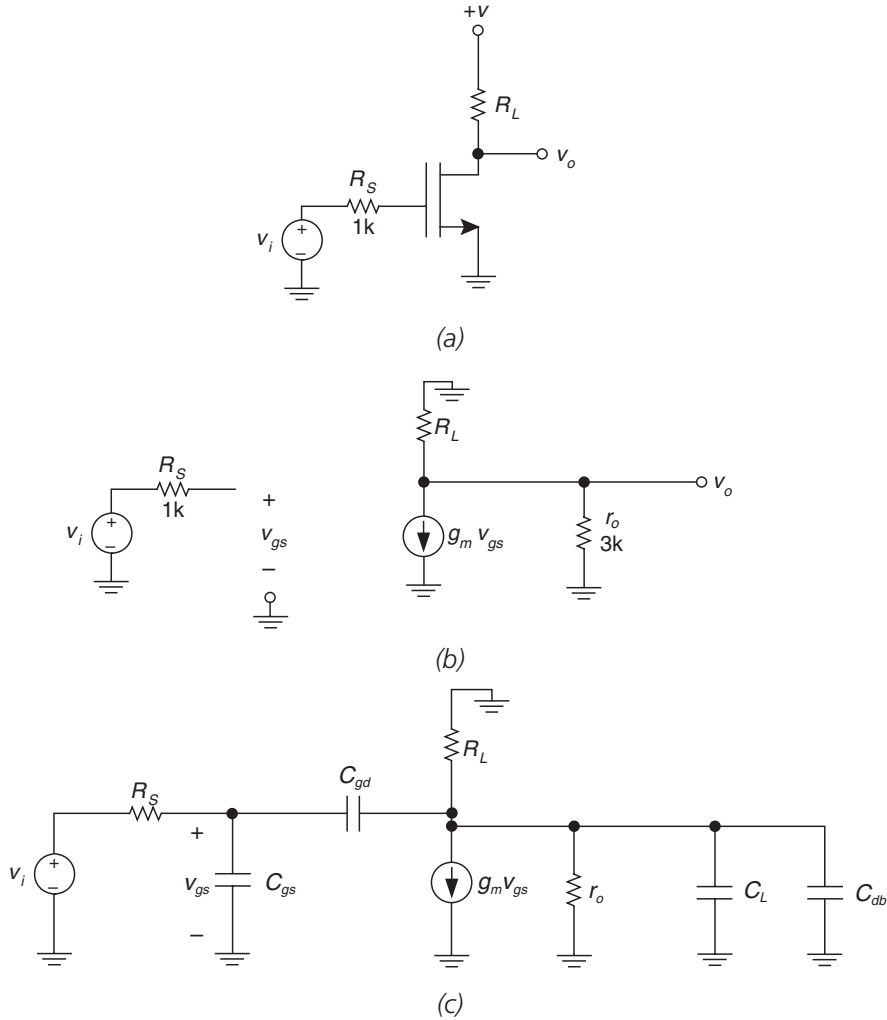


Figure 9-20: Try #1, MOS amplifier. (a) Amplifier, omitting biasing details for clarity. (b) Small-signal (low-frequency) model assuming $g_{mb} = 0$. (c) High-frequency small-signal model.

The circuits for finding the open-circuit time constants are shown in **Figure 9-21**. For the gate-source and gate-substrate capacitances, we use the circuit of **Figure 9-21a** to find:

$$R_{o1} = R_s = 1 \text{ k}\Omega$$

$$\tau_{o1} = R_{o1} C_{gs} = (1000)(0.2 \text{ pF}) = 0.2 \text{ ns} \quad [9-27]$$

For the load and drain-substrate capacitances, we use the circuit of **Figure 9-21b** to find:

$$R_{o2} = R_L \parallel r_o = 1043 \Omega$$

$$\tau_{o2} = R_{o2} (C_L + C_{db}) = (1043)(1.02 \text{ pF}) = 1.06 \text{ ns} \quad [9-28]$$

For the gate-drain capacitances, we use the circuit of **Figure 9-21c** to find:

$$\begin{aligned} R_{o3} &= R_S + R_L \parallel r_o + g_m R_S (R_L \parallel r_o) = 12473 \Omega \\ \tau_{o3} &= R_{o3} C_{gd} = (12473)(0.05 \text{ pF}) = 0.62 \text{ ns} \end{aligned} \quad [9-29]$$

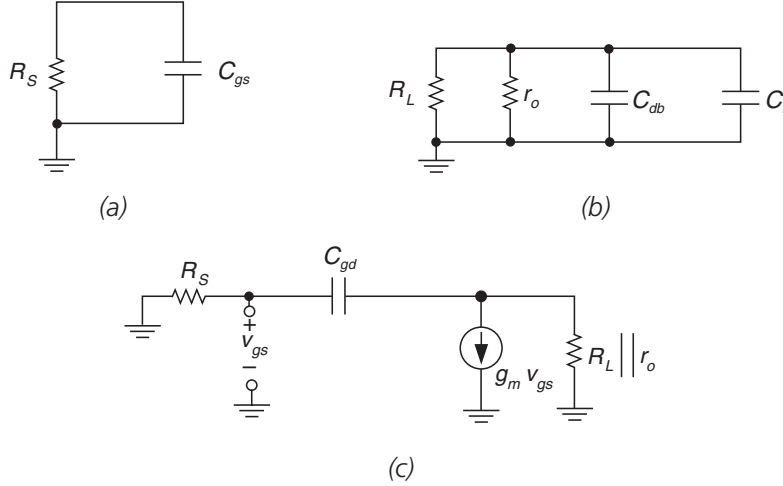


Figure 9-21: Circuits for finding open-circuit time constants for Try #1.

(a) Circuit for C_{gs} . (b) Circuit for $C_{db} + C_L$. (c) Circuit for C_{gd} .

The sum of open-circuit time constants and estimate of bandwidth are as follows:

$$\begin{aligned} \sum \tau_{oi} &= \tau_{o1} + \tau_{o2} + \tau_{o3} = 1.88 \text{ ns} \\ \omega_h &\approx \frac{1}{\sum \tau_{oi}} \approx \frac{1}{1.98 \times 10^{-9}} = 532 \times 10^6 \text{ rad/sec} \\ f_h &\approx \frac{\omega_h}{2\pi} \approx 84.7 \text{ MHz} \end{aligned} \quad [9-30]$$

We note that we won't meet the 250-MHz bandwidth specification yet, which is confirmed by a PSPICE simulation (**Figure 9-22**) which shows a bandwidth of ~92 MHz. The method of open-circuit time constants identifies the output load capacitance as the dominant bandwidth limitation. In order to improve the bandwidth, we'll isolate the load capacitance from the high-gain node by using a source follower.

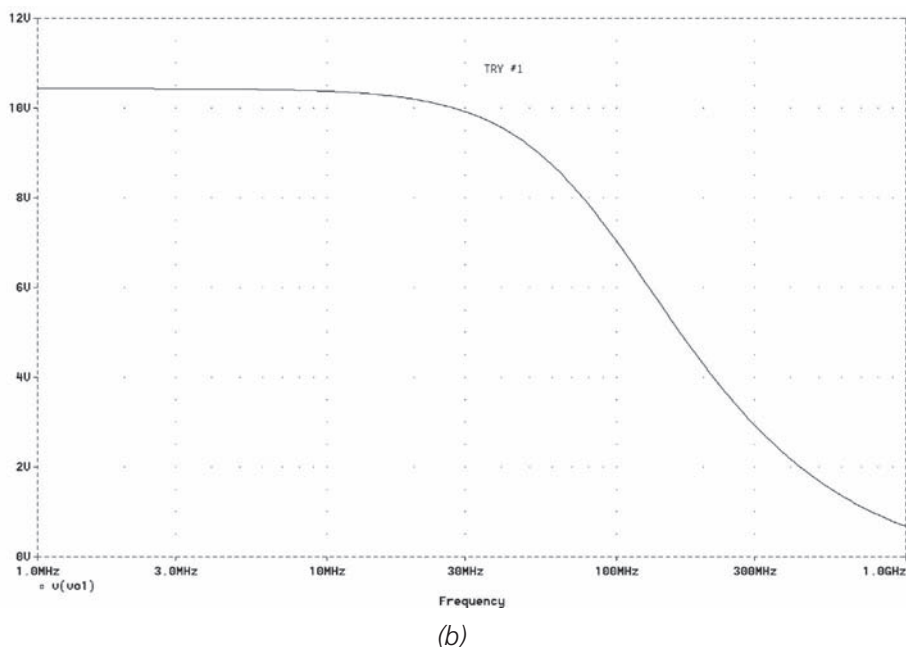
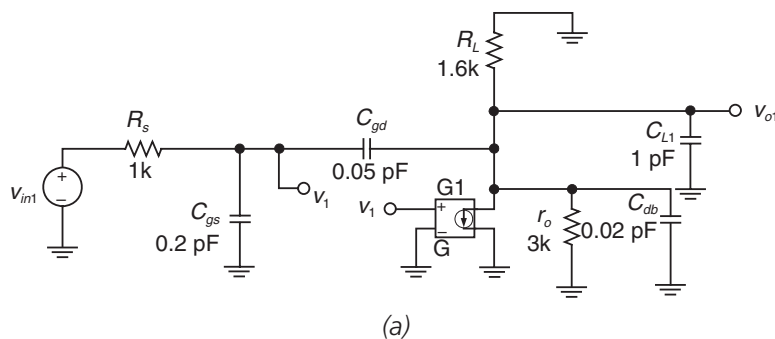


Figure 9-22: PSPICE result showing gain and bandwidth of MOS amplifier Try #1. (a) Circuit which was simulated. (b) PSPICE simulation result, showing a gain of -10.4 and a -3dB bandwidth of approximately 92 MHz .

TRY #2: Add output source follower

In Try #2, we've added device M_2 which is a source-follower buffer (**Figure 9-23**). We'll assume that the current source I_{BIAS2} biases M_2 at the same drain current as M_1 , and hence $g_{m2} = g_{m1}$. We'll also assume that all other MOS capacitances are the same for both devices.

Next, we perform an open-circuit time constants analysis using the small-signal model of **Figure 9-23b**.

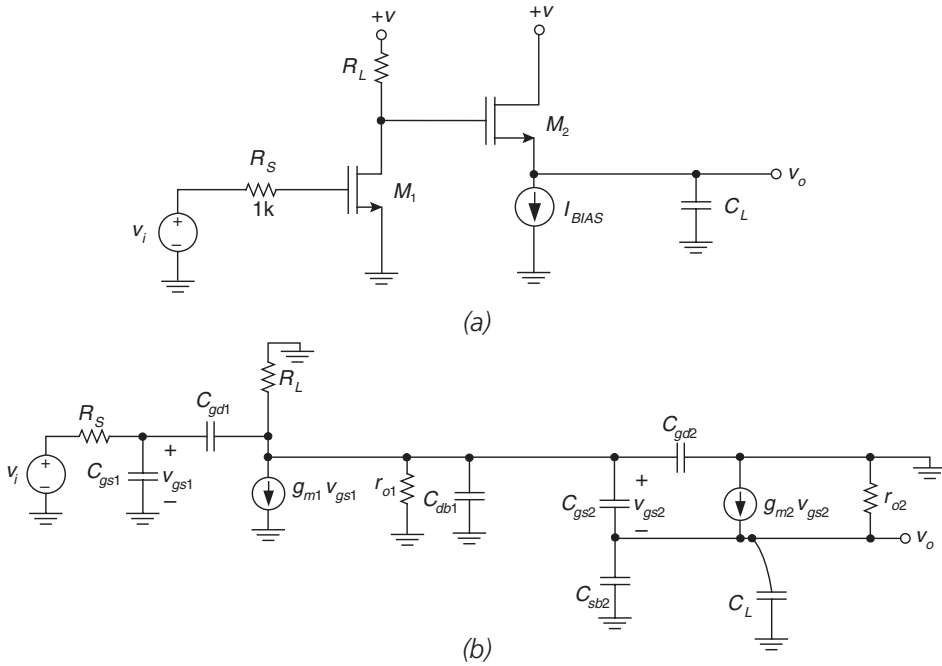


Figure 9-23: Common-source amplifier M_1 with source follower M_2 added.
(a) Circuit. (b) Small-signal model.

For M_1 , we note that the open-circuit time constants for C_{gs1} and for C_{gd1} are unchanged from the previous design iteration.⁹

$$\begin{aligned}\tau_{o1} &= 0.2 \text{ ns} \\ \tau_{o2} &= 0.62 \text{ ns}\end{aligned}\tag{9-31}$$

In the previous iteration, the high-gain node was loaded by $(C_L + C_{db})$. In this iteration, we have C_{db1} in parallel with C_{gd2} at the high-gain node as shown in **Figure 9-24a**. The open-circuit time constant for these capacitors are found as follows:

$$\begin{aligned}R_{o3} &= R_L \parallel r_{o1} = 1043\Omega \\ \tau_{o3} &= R_{o3} (C_{db1} + C_{gd2}) = (1043)(0.07 \text{ pF}) = 0.07 \text{ ns}\end{aligned}\tag{9-32}$$

For the gate-source capacitance of M_2 , we use the test circuit of **Figure 9-24b**, where we've added a test voltage source across the C_{gs2} terminals. By inspection,¹⁰ we find that the test current $i_t = g_{m2}v_t$, and hence the open-circuit time constant for C_{gs2} is as follows:

$$\begin{aligned}R_{o4} &= \frac{1}{g_{m2}} = 100\Omega \\ \tau_{o4} &= R_{o4}C_{gs2} = (100)(0.2 \text{ pF}) = 0.02 \text{ ns}\end{aligned}\tag{9-33}$$

⁹ Note that I'm renumbering the " τ_s " with this iteration.

¹⁰ I've ignored r_{o2} , which is OK in this case.

For the $C_L + C_{sb2}$ combination at the source terminal of M_2 , we use the test circuit of **Figure 9-24c**, where we've added a test voltage source across the source-follower output terminals. By inspection,¹¹ we find that the test current $i_t = g_{m2}v_t$, and hence the open-circuit time constant for $C_L + C_{sb2}$ is as follows:

$$R_{o5} = \frac{1}{g_{m2}} = 100\Omega$$

$$\tau_{o5} = R_{o5}(C_L + C_{sb2}) = (100)(1.02 \text{ pF}) = 0.1 \text{ ns}$$
[9-34]

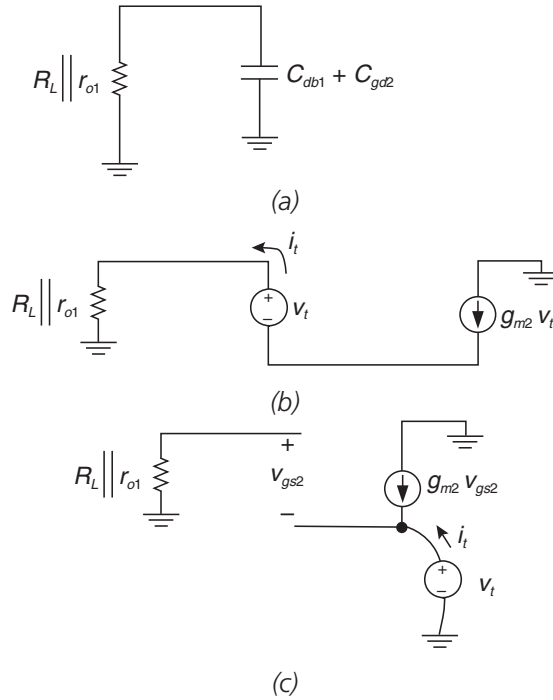


Figure 9-24: Open-circuit time constants circuits for Try #2.

(a) Circuit for finding open-circuit time constants for $C_{db1} + C_{gd2}$.

(b) Circuit for finding OCTCs for C_{gs2} . (c) Circuit for finding OCTCs for $C_L + C_{sb2}$.

A summary of the open-circuit time constants for Try #2 is shown in **Table 9-2**. We expect not to meet the bandwidth specification yet, since our open-circuit time constants estimate is only 157.6 MHz. PSPICE (**Figure 9-25**) confirms this; the simulated bandwidth is approximately 185 MHz.

¹¹ Again, ignoring r_{o2} .

Table 9-2: Summary of open-circuit time constant results for Try #2.

τ_{o1} (from C_{gs1})	0.20 ns
τ_{o2} (C_{gd1})	0.62 ns
τ_{o3} ($C_{db1} + C_{gd2}$)	0.07 ns
τ_{o4} (C_{gs2})	0.02 ns
τ_{o5} ($C_L + C_{sb2}$)	0.10 ns
Sum of OCTCs	1.01 ns
Estimate of ω_h	990 Mrad/sec
Estimate of f_h	157.6 MHz

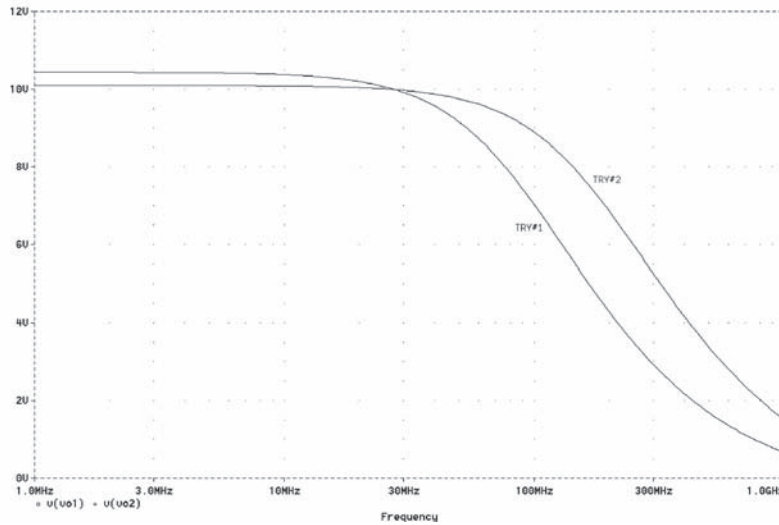
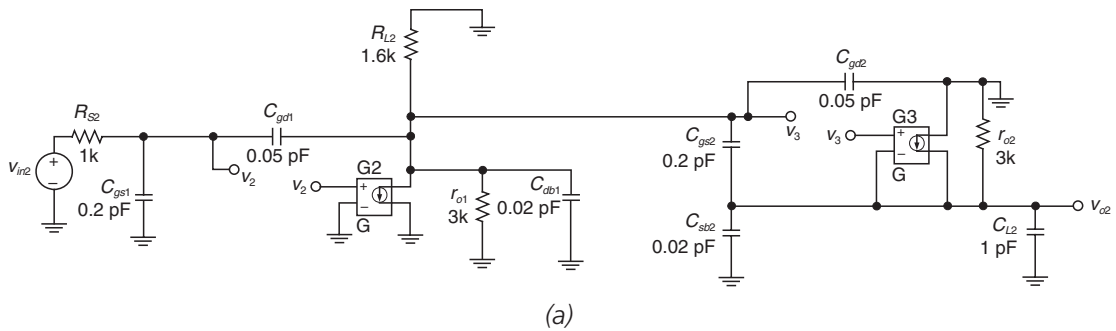


Figure 9-25: PSpice result showing gain and bandwidth of MOS amplifier Try #1 and Try #2. (a) Circuit. (b) PSpice simulation, showing comparison of Try #1 and Try #2. Try #2 has a higher bandwidth (~185 MHz) and a slightly lower gain (~10.1)

TRY #3: Add cascode transistor

In the last open-circuit time constants calculation, we found that the dominant time constant is due to the Miller effect of device M_1 . We will add a common-gate (cascode) transistor M_3 in order to try to reduce the Miller effect, resulting in the circuit of **Figure 9-26a**.

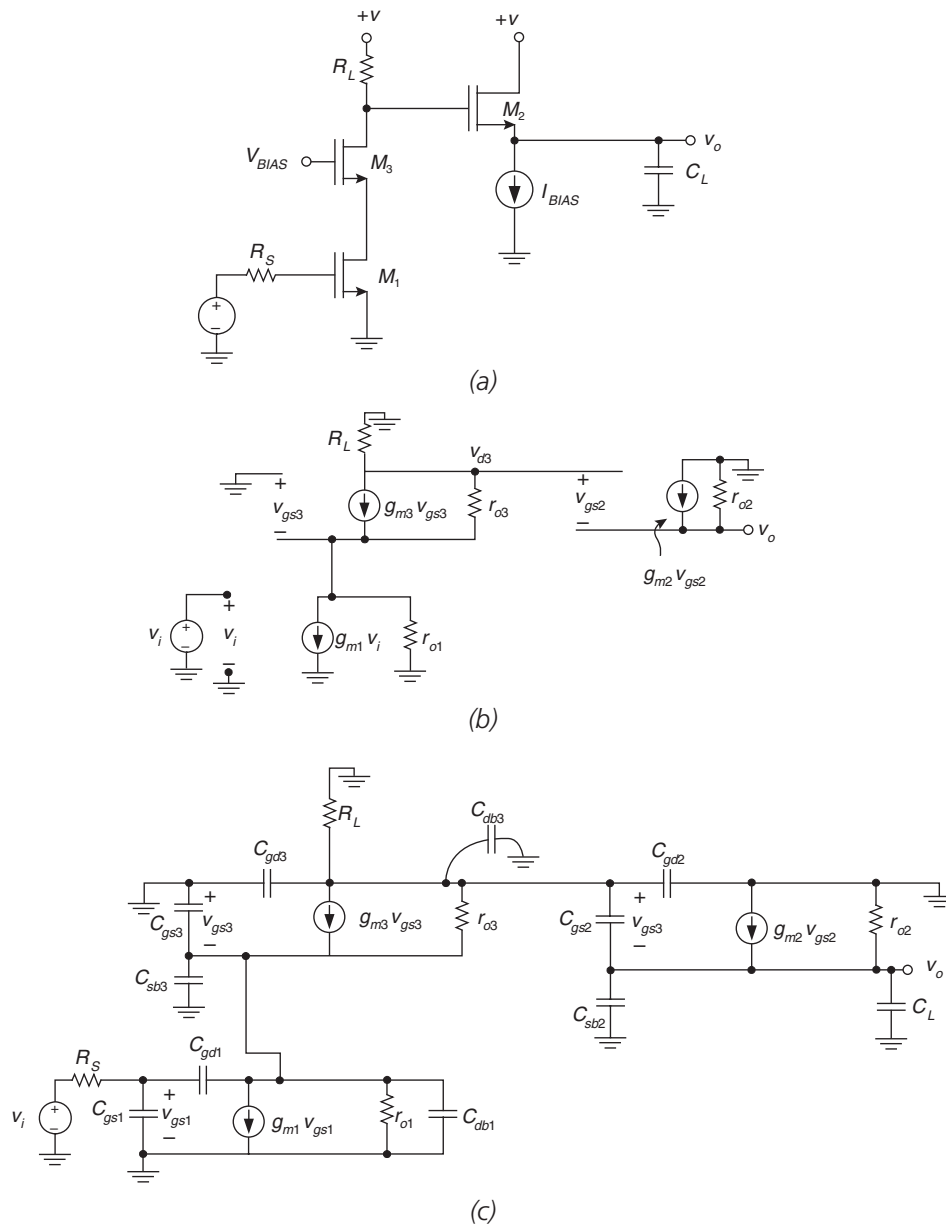


Figure 9-26: Try #3, with cascode transistor M_3 added.
(a) Circuit. (b) Low-frequency small-signal model. (c) High-frequency small-signal model.

The low-frequency small-signal model is shown in **Figure 9-26b**. We note from this model that the gain of this amplifier configuration is now roughly:

$$\frac{v_o}{v_i} \approx -g_{m1}R_L \quad [9-35]$$

This assumes that the loading effect of r_{o3} is minimized at the high-gain node and that the gain of the M_3 emitter-follower is close to unity. We can argue that r_{o3} doesn't significantly affect the gain by noting that the incremental output resistance seen at the drain of M_3 is increased due to the fact that the source of M_3 is loaded with a relatively high resistance (i.e., r_{o1}). Thus, the effects of loading due to the output resistance r_{o3} is minimized. We can then reduce the load resistor in this iteration to perhaps 1.1 k Ω (reduced from 1.6 k Ω) and still meet the gain specification. This will also help us meet the bandwidth specification.

Referring to the small-signal circuit of **Figure 9-26c**, we now need to do some bookkeeping to keep track of all the various capacitances and open-circuit time constant calculations. There are 13 capacitances that contribute to open-circuit time constants: four for each transistor plus the load capacitance. However, we find that some capacitances are shorted out since all transistors have the bulk terminal grounded (i.e., C_{sb1} and C_{db2}) and that there are many capacitances that are in parallel (i.e., $C_{db1} + C_{gs3} + C_{sb3}$; $C_{sb2} + C_L$; and $C_{gd2} + C_{gd3} + C_{db3}$). The open-circuit time constant calculations for the various capacitances are shown in **Table 9-2**.

Table 9-2: Summary of OCTC calculations for Try #3.

Capacitance(s)	Open-circuit resistance	Open-circuit time constant calculation
$C_{gs1} = 0.2 \text{ pF}$	Unchanged from previous iteration at 1 k Ω	τ_{o1} unchanged at 0.2 ns
$C_{gd1} = 0.05 \text{ pF}$	Open-circuit resistance is greatly reduced due to low output resistance $\approx r_{out3} = 100\Omega$ at source of M_3 . $R_{o2} = R_s + r_{out3} + g_{m1}R_s r_{out3} = 2100\Omega$	$\tau_{o2} = (2100)(0.05 \text{ pF}) = 0.105 \text{ ns}$
$C_{db1} + C_{sb3} + C_{gs3} = 0.24 \text{ pF}$	Open-circuit resistance $= r_{out3} \approx 1/g_{m3} = 100\Omega$	$\tau_{o3} = (100)(0.24 \text{ pF}) = 0.024 \text{ ns}$
$C_{gs2} = 0.2 \text{ pF}$	Open-circuit resistance unchanged at $1/g_{m2} = 100\Omega$	$\tau_{o4} = (100)(0.2 \text{ pF}) = 0.02 \text{ ns}$
$C_{gd2} + C_{db3} + C_{gd3} = 0.12 \text{ pF}$	Resistance from this common node to ground $\approx R_L = 1100\Omega$	$\tau_{o5} = (1100)(0.12 \text{ pF}) = 0.132 \text{ ns}$
$C_{sb2} + C_L = 1.02 \text{ pF}$	Open-circuit resistance $= 1/g_{m2} = 100\Omega$	$\tau_{o5} = (100)(1.02 \text{ pF}) = 0.10 \text{ ns}$

The sum of the open-circuit time constants is now 0.58 nanoseconds, resulting in a bandwidth estimate of 1721 Mrad/seconds, or 274 MHz. This circuit looks like a good candidate for final simulation, since we know the method of open-circuit time constants is always conservative in estimating bandwidth. A detailed PSPICE simulation shows that the bandwidth is 460 MHz (**Figure 9-27**), and hence we meet our 350 MHz bandwidth specification.

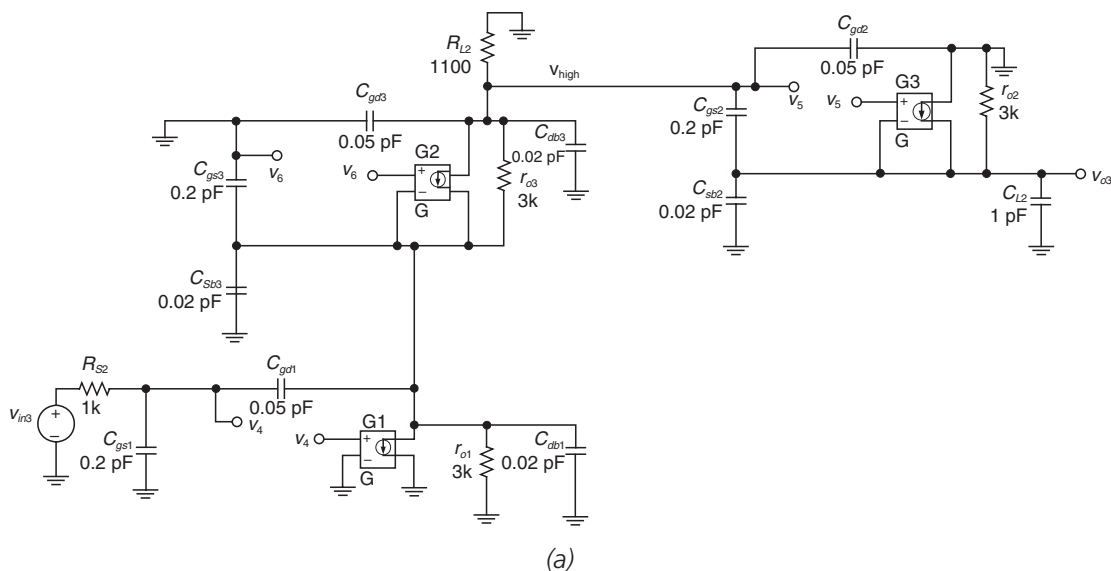


Figure 9-27: PSPICE result showing gain and bandwidth of MOS amplifier Try #3.

(a) Circuit, with input common-source amplifier (M_1), cascode transistor (M_2) and output source follower (M_3). (b) PSPICE result, showing low-frequency gain of -10.2 and a -3 dB bandwidth of 460 MHz.

If further bandwidth enhancement was needed, we could add an input source-follower to isolate the source resistance from the input capacitances of M_1 , as shown in **Figure 9-28a**. This attacks the largest open-circuit time constant which is associated with C_{gs1} . Adding the

emitter-follower increases the bandwidth to over 800 MHz with some gain peaking, as shown in **Figure 9-28b**.

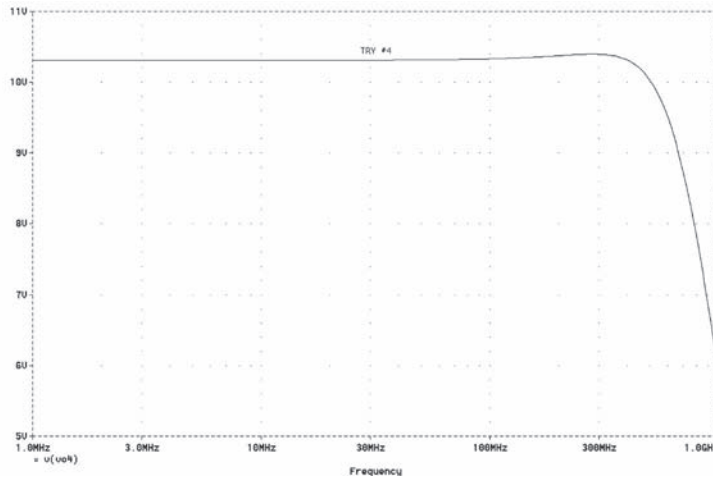
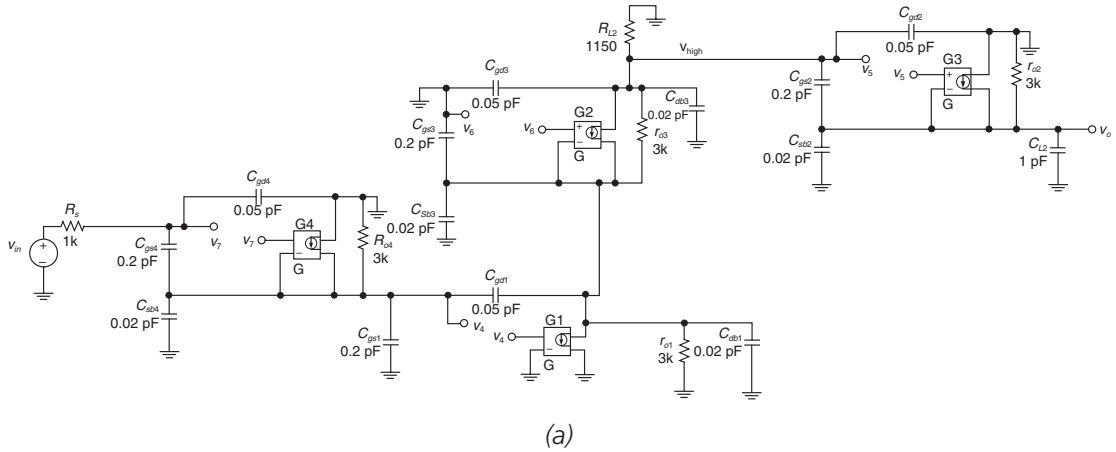


Figure 9-28: PSPICE result showing gain and bandwidth of MOS amplifier Try #4 where an input source follower was added to isolate M_1 's input capacitances from the $1\text{-k}\Omega$ source resistance. (a) Circuit. Note that load resistor was slightly increased to 1150 ohms to meet the gain specification. (b) PSPICE result, showing low-frequency gain of -10.3 and a -3dB bandwidth of 855 MHz.

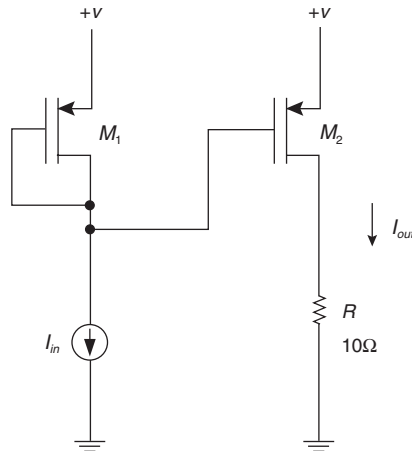
Chapter 9 Problems

Problem 9.1

A MOS current mirror is shown **Figure 9-29**. Assume that each transistor has $g_m = 0.01$ A/V, $C_{gs} = 1$ pF and $C_{gd} = 0.5$ pF and that all other transistor parasitics are negligible.

- Draw the small-signal incremental model of the circuit.
- Estimate the bandwidth of the circuit using open-circuit time constants.

Figure 9-29: MOS current mirror for Problem 9.1.

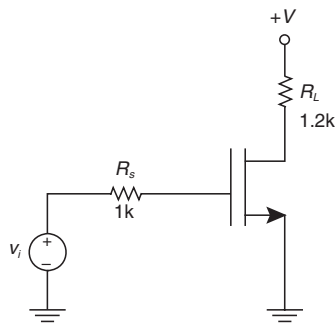


Problem 9.2

A MOS common-source amplifier is shown in **Figure 9-30**. Assume that the transistor has $g_m = 0.02$ A/V, $C_{gs} = 1$ pF and $C_{gd} = 0.5$ pF and transistor incremental output resistance $r_o = 5$ k Ω .

- Draw the small-signal incremental model of the circuit.
- Find the low-frequency gain.
- Estimate the bandwidth of the circuit using open-circuit time constants.

Figure 9-30: MOS common-source amplifier for Problem 9.2, omitting biasing details.



Problem 9.3

Using the four-transistor MOS amplifier (**Figure 9-28**) from the CMOS amplifier design example in this chapter as a basis, suggest possible circuit modifications that would further improve bandwidth. Verify your design modifications using open-circuit time constants and PSPICE simulations.

References

- Arns, R. G., "The other transistor: early history of the metal-oxide semiconductor field-effect transistor," *Engineering Science and Education Journal*, vol. 7, no. 5, October 1998, pp. 233–240.
- Comer, David J., and Comer, Donald T., "Teaching MOS Integrated Circuit Amplifier Design to Undergraduates," *IEEE Transactions on Education*, vol. 44, no. 3, August 2001, pp. 232–238.
- Gray, P. R., and Meyer, R. G., "MOS operational amplifier design-a tutorial overview," *IEEE Journal of Solid State Circuits*, vol. 17, no. 6, December 1982, pp. 969–982.
- Grebene, Alan B., *Bipolar and MOS Analog Integrated Circuits*, John Wiley, 1984.
- Hastings, Alan, *The Art of Analog Layout*, Prentice Hall, 2001.
- Hodges, D. A., Gray, P. R., and Brodersen, R. W., "Potential of MOS technologies for analog integrated circuits," *IEEE Journal of Solid State Circuits*, vol. 13, no. 3, June 1978, pp. 285–294.
- Jiang, Ruoxin, Tang, Haiming, and Mayaram, Kartikeya, "A Simple and Accurate Method for Calculating the Low Frequency Common-Mode Gain in a MOS Differential Amplifier with a Current-Mirror Load," *IEEE Transactions on Education*, vol. 43, no. 3, August 2000, pp. 362–364.
- Johns, David A., and Martin, Ken, *Analog Integrated Circuit Design*, John Wiley, 1997.
- Lee, Thomas H., *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge University Press, 1998.
- Lilienfeld J., US Patents 1,745, 175; 1,877, 140 and 1,900,018, available from the U.S. Patent and Trademark office, www.uspto.gov.
- Middlebrook, R. D., and Richer, I., "Limits on the Power-Law Exponent for Field-Effect Transistor Transfer Characteristics," *Solid-State Electronics*, vol. 6, September-October 1963, pp. 542–544.
- Pierret, Robert F., *Modular Series on Solid State Devices*, Field Effect Devices, (volume 4), Addison-Wesley, 1983. Singh, Jasprit, *Semiconductor Devices Basic Principles*, John Wiley, 2001.
- Sze, S. M., *Semiconductor Devices Physics and Technology*, 2nd edition, John Wiley, 2002.
- Winarski, T. Y., "Dielectrics in MOS devices, DRAM capacitors, and inter-metal isolation," *IEEE Electrical Insulation Magazine*, vol. 17, no. 6, November-December 2001, pp. 34–47.

Bipolar Transistor Switching and the Charge Control Model

In This Chapter

- In this chapter, we examine in detail large-signal switching of bipolar transistors. A model used to the switching speed, called the charge control model, is introduced.

Introduction

In previous chapters, we discussed the dynamic behavior of transistors from a small-signal point of view using the hybrid- π model. In other words, the small-signal models developed so far are valid only for small variations of the voltages and currents around an operating point. We used the method of open-circuit time constants to analyze bandwidth limitations in linear amplifiers.

A further modeling technique, called the *charge control model*, permits us to analyze the behavior of transistor circuits during large-signal switching and during transistor saturation. Similarly to the method of open-circuit time constants, the charge control model gives results that aren't necessarily numerically accurate; rather the usefulness of the method is the design insight that we derive from the application of relatively simple models.

Development of the Switching Models

In the following development, we consider the charge control model for the NPN transistor. The results are exactly the same for a PNP transistor. Transistor terminal current definitions are given in **Figure 10-1**.

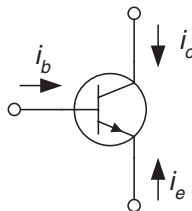
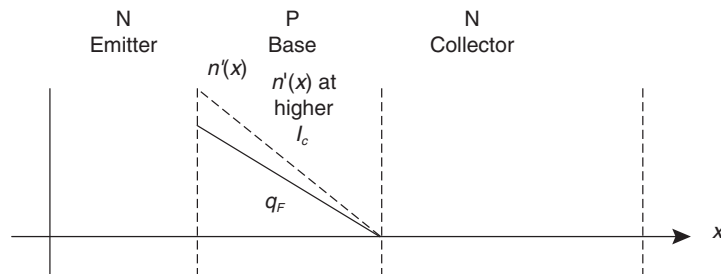


Figure 10-1: Terminal definitions for NPN transistor, showing base current i_b , collector current i_c and emitter current i_e .

From device physics, we derived in an earlier chapter the fact that when a transistor is biased in the forward-active region the base-to-emitter junction is forward biased and there is extra charge stored in the base region. For an NPN transistor, the *excess minority carrier concentration* is shown in **Figure 10-2**. In the case of the NPN transistor, the base is made of P material and the excess minority carriers are electrons, and have the profile $n(x)$ shown. There is a similar hole profile in the emitter $p(x)$ but for purposes of analyzing switching operation we need only consider the base region.

Since transistor collector current is dominated by diffusion in the base, collector current I_c is proportional to the slope of the $n'(x)$ curve. The dotted line shows the carrier concentration profile at a higher collector current. Since the vertical axis has units of charge concentration, the total area under the curve is proportional to the charge stored in the base of the transistor. When the transistor is operated in the normal fashion and in the linear region (called the forward-active region) we'll call this stored charge q_F , for *forward charge*.

Figure 10-2: Excess carrier concentration for NPN transistor in forward-active region. The extra charge in the base ($n'(x)$) causes a net diffusion current of electrons to the right, resulting in a collector current to the left.



We can summarize this relationship between collector current and the charge stored in the base as:

$$i_c = \frac{q_F}{\tau_F} \quad [10-1]$$

where τ_F is a device-dependent constant that has units of time. This is the *charge-control equation* for the collector current in the forward-active region. We assume that as the base-emitter voltage changes, the concentration $n'(x=0)$ changes instantaneously in response.¹

Now, we need to consider the base current. Since there is charge stored in the base, we expect that the base current charge control equation has a term related to how fast we add and remove charge q_F from the base. Furthermore, we know that there has to be a DC term (because collector and base current bias levels are related by the large signal β_F). Using this reasoning, the base current charge control equation is:

$$i_b = \frac{q_F}{\tau_{BF}} + \frac{dq_F}{dt} \quad [10-2]$$

¹ This is approximately true as long as we don't allow the transistor v_{be} to change too fast.

The q_F/τ_{BF} term is the DC base current, and the dq_F/dt term is the charge needed if we change the charge stored in the base. The characteristic time τ_{BF} is related to τ_F in a way that we'll explain later.

We also recognize that the emitter current is the negative sum of the collector and base currents, or:

$$i_e = -(i_c + i_b) \quad [10-3]$$

Summarizing, the charge control equations for the forward-active region are:

$$\begin{aligned} i_b &= \frac{q_F}{\tau_{BF}} + \frac{dq_F}{dt} \\ i_c &= \frac{q_F}{\tau_F} \\ i_e &= -(i_c + i_b) \end{aligned} \quad [10-4]$$

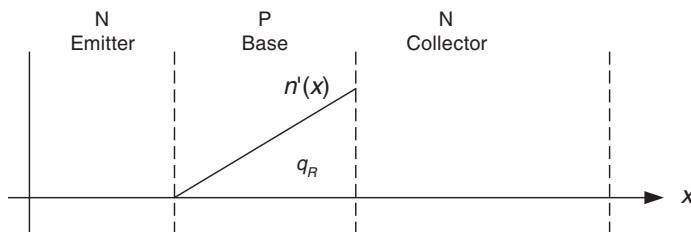
Note that these equations do not account for charge stored in the base-emitter and base-collector junction capacitances. We'll fix this later.

Reverse-Active Region

In the forward-active region, the base-emitter junction is forward biased and the collector-base junction is reverse biased. We could just as easily operate the transistor with the collector-base junction forward biased, and the base-emitter junction reverse biased. This mode of operation is called the *reverse-active region*, and analysis is exactly analogous with the forward-active region, with the roles of the collector and emitter reversed.² The carrier profile is as shown in **Figure 10-3**. The charge control equations in the reverse-active region are:

$$\begin{aligned} i_b &= \frac{q_R}{\tau_{BR}} + \frac{dq_R}{dt} \\ i_e &= \frac{q_R}{\tau_R} \\ i_c &= -(i_e + i_b) \end{aligned} \quad [10-5]$$

Figure 10-3: Excess carrier concentration for NPN transistor in reverse-active region. Note that in the reverse-active region the collector-base junction is forward biased and the base-emitter junction is reverse biased.



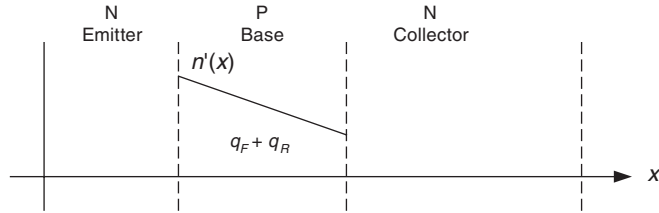
² The charge stored in the base in the reverse-active region is q_R .

Transistors are optimized to run in the forward active region. When operated in the reverse active region, as might be expected, transistor parameters such as speed and β_F are degraded.

Saturation

Saturation is the case when both base-emitter and base-collector junctions are forward biased. We can consider this mode of operation as a mixture of the forward-active and reverse-active operation, as shown in **Figure 10-4**.

Figure 10-4: Excess carrier concentration for NPN transistor in saturation region. In saturation, both junctions are forward biased.



The total charge control equations including forward-active and reverse-active operation are found by summing the previous results:

$$\begin{aligned} i_b &= \frac{q_F}{\tau_{BF}} + \frac{dq_F}{dt} + \frac{q_R}{\tau_{BR}} + \frac{dq_R}{dt} \\ i_c &= \frac{q_F}{\tau_F} - \frac{dq_R}{dt} - q_R \left[\frac{1}{\tau_R} + \frac{1}{\tau_{BR}} \right] \\ i_e &= -(i_c + i_b) \end{aligned} \quad [10-6]$$

Since we've shown that the forward active region is dominated by one time constant (τ_F) and the reverse active region is dominated by another time constant (τ_R), it makes sense that the saturation region will have two time constants, but not necessarily τ_F and τ_R . In fact, we'll show that the dynamics in saturation has time constants that are some weighted functions of these individual time constants.

We can solve for the natural frequencies of the growth of charge in the transistor by solving the homogeneous case (and making the Laplace substitution $d/dt \rightarrow s$):

$$\begin{aligned} 0 &= q_F \left(s + \frac{1}{\tau_{BF}} \right) + q_R \left(s + \frac{1}{\tau_{BR}} \right) \\ 0 &= \frac{q_F}{\tau_R} - q_R \left(s + \frac{1}{\tau_R} + \frac{1}{\tau_{BR}} \right) \end{aligned} \quad [10-7]$$

We can solve these simultaneous equations, resulting in:

$$s^2 + \left(\frac{1}{\tau_F} + \frac{1}{\tau_{BF}} + \frac{1}{\tau_R} + \frac{1}{\tau_{BR}} \right) s + \left(\frac{1}{\tau_F \tau_{BR}} + \frac{1}{\tau_R \tau_{BF}} + \frac{1}{\tau_{BR} \tau_{BF}} \right) = 0 \quad [10-8]$$

or, simplifying:⁴

$$s^2 + \left(\frac{\beta_F + 1}{\tau_{BF}} + \frac{\beta_R + 1}{\tau_{BR}} \right) s + \left(\frac{1}{\tau_F \tau_{BR}} + \frac{1}{\tau_R \tau_{BF}} + \frac{1}{\tau_{BR} \tau_{BF}} \right) = 0 \quad [10-9]$$

In general, there are two widely spaced poles,³ with the “fast” pole being comparable to the ω_T of the transistor:

$$s_{fast} \approx \frac{\beta_F + 1}{\tau_{BF}} + \frac{\beta_R + 1}{\tau_{BR}} \approx \frac{1}{\tau_F} + \frac{1}{\tau_R} = \omega_T \quad [10-10]$$

The low-frequency pole is given by:

$$s_{slow} \approx \frac{\beta_F + \beta_R + 1}{\tau_{BF}(\beta_R + 1) + \tau_{BR}(\beta_F + 1)} \quad [10-11]$$

The fast time constant, sometimes called the “slosh” mode, corresponds to the time scale with which charge redistributes between q_F and q_R . The slow time constant, or the “fill” mode, corresponds to how fast q_F and q_R rise together. This is analogous to even and odd modes in oscillating mass-spring systems. The slosh mode dies out with a very fast time constant, and therefore the charge growth in the saturation region is dominated by the slow time constant, as:

$$\tau_s = \frac{\tau_{BF}(\beta_R + 1) + \tau_{BR}(\beta_F + 1)}{\beta_F + \beta_R + 1} \quad [10-12]$$

Hence, we can approximate the charge control equation in the saturation region as:

$$i_b - \frac{I_{C,SAT}}{\beta_F} = \frac{q_s}{\tau_s} + \frac{dq_s}{dt} \quad [10-13]$$

The term τ_s is the saturation time constant, and $I_{C,SAT}$ is the collector current in saturation. The term q_s is the saturation charge.

³ Remember our widely spaced pole approximation. Let’s assume that you have a 2-pole transfer function of the form:

$$H(s) = \frac{1}{s^2 + As + B}$$

If the poles are on the real axis and are widely spaced, we can approximate the pole locations by:

$$s_{fast} \approx -A$$

$$s_{slow} \approx -B/A$$

Let’s try this method on a transfer function

$$H(s) = \frac{1}{s^2 + 11s + 10} = \frac{1}{(s+1)(s+10)}$$

There is a fast pole at -10 radians/second and a slow dominant pole at -1 radian/second. If we apply the widely spaced pole approximation, we get $s_{fast} \approx -11$ and $s_{slow} \approx -10/11$.

⁴ We haven’t worked it out yet, but we use here the identities $\tau_{\beta F} = \beta_F \tau_F$ and $\tau_{\beta R} = \beta_R \tau_R$ where β_F is forward beta and β_R is reverse beta. More on this identify later on.

Junction Capacitances

In addition to the stored charge in the base region, there is charge stored in the nonlinear base-emitter and base-collector junction capacitances. We'll term these stored charges q_{ve} and q_{vc} , and this corresponds to the large-signal version of the base capacitance C_{je} and feedback capacitance C_{jc} . The extra terms that we need to add to the previous charge control equations are:

$$\begin{aligned} i_{b,SCL} &= \frac{d}{dt}(q_{ve} + q_{vc}) \\ i_{c,SCL} &= -\frac{d}{dt}q_{vc} \end{aligned} \quad [10-14]$$

The notation "SCL" refers to the fact that these current components of base and collector currents charge and discharge the *space charge layers*, also called the *depletion* capacitances. We'll see how to find these capacitances on a datasheet later on.

Relationship Between Charge Control and Hybrid-Pi Parameters

We'll next work out an important relationship between charge control parameters and hybrid-pi parameters. In **Figure 10-5a** we see a circuit that will help us work out this relationship. The small-signal hybrid-pi model for this circuit is shown in **Figure 10-5b**.

If we run the transistor at very high frequencies, the hybrid-pi model finds the base current:

$$i_b|_{\omega \rightarrow \infty} \approx (C_\pi + C_\mu) \frac{dv_{be}}{dt} \quad [10-15]$$

For very high frequencies, the charge control model predicts (ignoring space-charge capacitances):

$$i_b \approx \frac{dq_F}{dt} \approx \frac{d}{dt}(i_c \tau_F) \approx g_m \tau_F \frac{d}{dt}(v_{be}) \quad [10-16]$$

We find the following relationship:

$$\tau_F \approx \frac{(C_\pi + C_\mu)}{g_m} \approx \frac{1}{\omega_T} \quad [10-17]$$

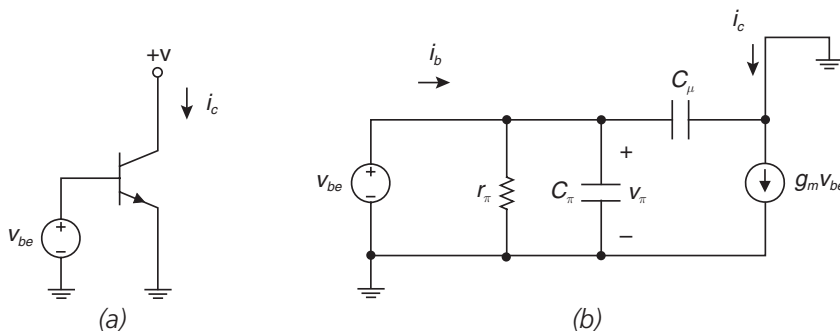


Figure 10-5: Circuit for finding relationship between charge control and hybrid-pi parameters. (a) Transistor circuit. (b) High-frequency hybrid-pi model.

Next, by looking at the charge control equations at very low frequencies, we find:

$$\begin{aligned}
 i_b|_{DC} &= \frac{q_F}{\tau_{BF}} \\
 i_c &= \frac{q_F}{\tau_F} \\
 \frac{i_c}{i_b|_{DC}} &= \frac{\tau_{BF}}{\tau_F} \rightarrow \tau_{BF} = \beta_F \tau_F
 \end{aligned}
 \tag{10-18}$$

So, we now have a relationship between the charge control parameters τ_F and τ_{BF} and the hybrid-pi parameters β_F and ω_T .

Finding Junction Capacitances from the Datasheet

How do we find the values of junction capacitances? Well, we're fortunate if manufacturers' data is available. For instance, for the 2N3904 transistor, capacitance data is available for reverse bias voltages in excess of 0.1 volts (i.e., junction voltages less than $-0.1V$) on both the emitter-base and collector-base junctions, as shown in **Figure 10-6**.

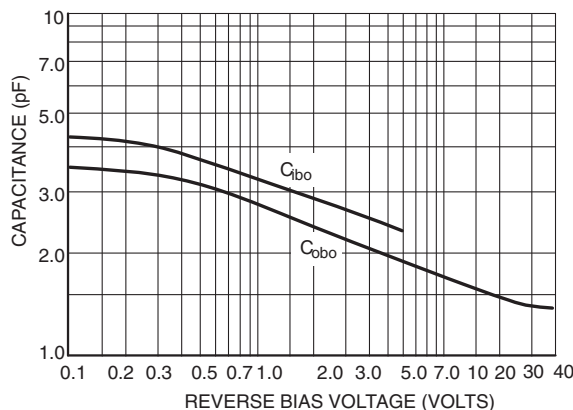


Figure 10-6: Manufacturer's datasheet⁵ tabulation of junction capacitance of 2N3904 transistor.

What do we do if we need charge data for *positive* voltages on the base-emitter junction? An NPN transistor is cutoff for V_{BE} less than 0.4V or so. The datasheet doesn't give us capacitance values for positive V_{BE} . Fortunately, there is a little bit of physics that's available to us. The junction capacitance of a PN junction is given by:

$$C_j = \frac{C_{jo}}{\left[1 - \frac{V_j}{V_{BI}}\right]^m}
 \tag{10-19}$$

⁵ From On Semiconductor, <http://www.onsemi.com>, reprinted with permission of On Semiconductor.

C_{j0} is the junction capacitance at a junction voltage $V_j = 0$; V_{BI} is the *built-in voltage* characteristic of the junction and is given by:

$$V_{BI} = \frac{kT}{q} \ln \left[\frac{N_A N_D}{n_i^2} \right] \quad [10-20]$$

where N_A and N_D are the doping levels on the P and N sides, and n_i is the intrinsic minority carrier concentration. A typical value for V_{BI} is 0.8V at normal doping levels. The factor m depends on the type of junction, and is 0.5 for an abrupt junction and 0.333 for a linearly graded junction.

A curve fit for the emitter-base junction capacitance for the 2N3904 transistor is given in **Figure 10-7**, with $C_{j0} = 4.5$ picofarads, $V_{BI} = 0.8V$ and $m = 0.333$. Note that the junction capacitance blows up when the junction voltage approaches the built-in voltage. This curve fit will be sufficient for crude estimates of switching time using the charge-control model.

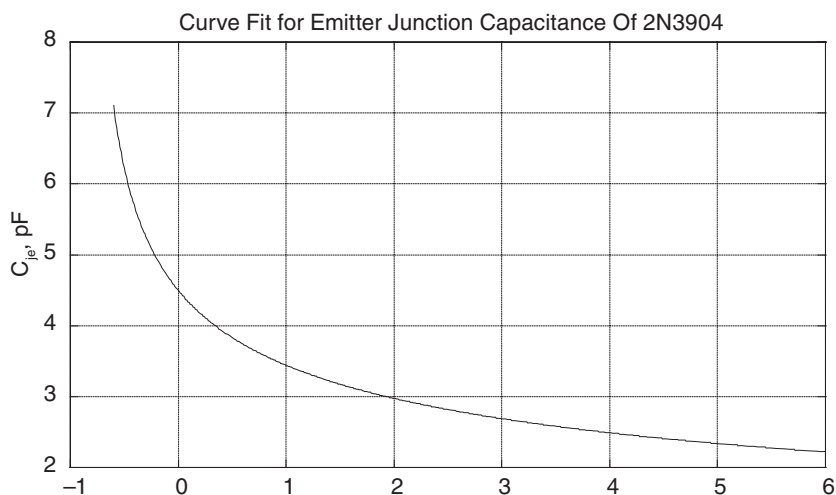
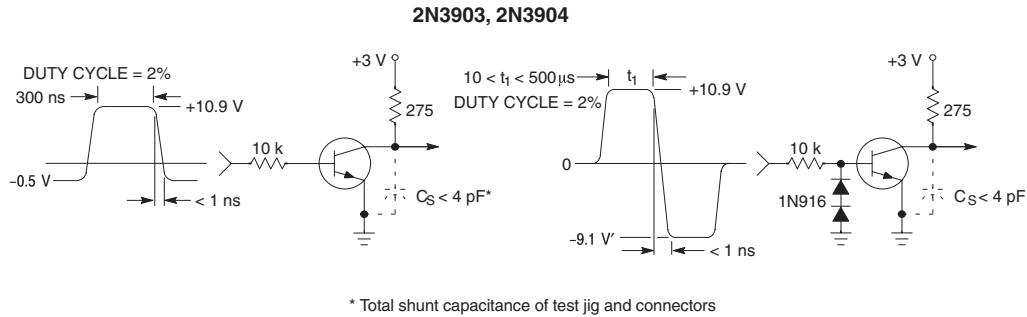


Figure 10-7: Curve fit for emitter-base junction capacitance C_{je} for 2N3904 transistor with $C_{j0} = 4.5$ picofarads, $V_{BI} = 0.8V$ and $m = 0.333$.

Manufacturers' Testing

Manufacturers sometimes (but not always) put switching time test results on their datasheets. Shown below are test circuits and test results for the 2N3904 NPN transistor.⁶ From these test results we can approximate various charge control parameters.



**Figure 1. Delay and Rise Time
Equivalent Test Circuit**

**Figure 2. Storage and Fall Time
Equivalent Test Circuit**

Figure 10-8: Manufacturer's test circuits for the 2N3904 transistor.

Referring to **Figure 10-9**, the delay time listed ($t_d = 35$ nanoseconds) is the delay before the transistor turns on. The rise time (t_r) is the 10–90% risetime of the collector current. Storage time⁷ (t_s) is the time it takes to remove saturation charge before coming out of saturation. The fall time t_f is the time it takes collector current to fall from 90% of full value to 10%.

SWITCHING CHARACTERISTICS				
Delay Time	$(V_{CC} = 3.0 \text{ Vdc}, V_{BE} = 0.5 \text{ Vdc}, I_C = 10 \text{ mAdc}, I_{B1} = 1.0 \text{ mAdc})$	t_d	–	35 ns
Rise Time		t_r	–	35 ns
Storage Time	$(V_{CC} = 3.0 \text{ Vdc}, I_C = 10 \text{ mAdc}, I_{B1} = I_{B2} = 1.0 \text{ mAdc})$	t_s	–	175 ns
Fall Time		t_f	–	50 ns

Figure 10-9: Manufacturer's switching time results found on 2N3904 transistor datasheet.

Charge Control Model Examples

Next, we'll see how to use the charge control model by working through a number of examples.

Example 10.1: Transistor inverter with base current drive

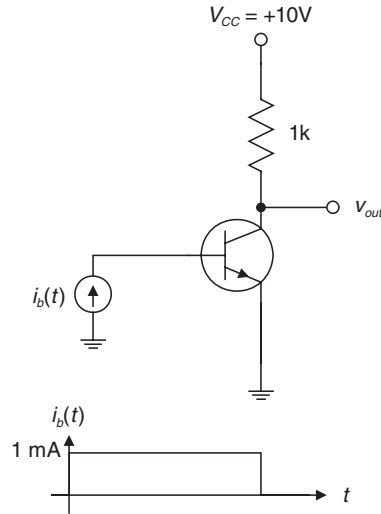
In this problem, a transistor is controlled by driving a base current as in **Figure 10-10**. Let's analyze the dynamics of the transistor, assuming transistor parameters: $\tau_F = 0.3$ nanoseconds, $\beta_F = 416$, $\tau_R = 240$ nanoseconds, and $\beta_R = 0.7$. We'll ignore space charge capacitances. We'll

⁶ Taken from On Semiconductor 2N3904 datasheet, found at <http://www.onsemi.com>, reprinted with permission of On Semiconductor.

⁷ Note that this storage delay time t_s is *not* the same as the saturation region time constant τ_s . We can use the value of t_s supplied by the manufacturer to estimate τ_s , however.

find the switching profile of this transistor circuit under these conditions. Assume that the current pulse transitions high at $t = 0$. Much later, after all transients have died down, the current pulse transitions back to zero.

Figure 10-10: Transistor with base current drive.



Crossing the forward-active region

Ignoring space charge capacitances, we'll find base charge $q_F(t)$ and collector current $i_c(t)$ when crossing through the forward-active region. Since we're ignoring space charge capacitances, the necessary charge control equation is relatively simple. First, recognize that when we first step the base current, the transistor will enter the forward active region. There's no turn-on delay time in this example since we have assumed that junction capacitances are negligible. After a while, the transistor will saturate, since

$$\beta_F I_B R_L > V_{CC} \quad [10-21]$$

When the transistor is in the forward-active region, the charge control equations are:

$$i_b = \frac{q_F}{\tau_{BF}} + \frac{dq_F}{dt} \quad [10-22]$$

$$i_c \approx \frac{q_F}{\tau_F}$$

Since we are driving the base with a stepped current source, the differential equation for the base current is:

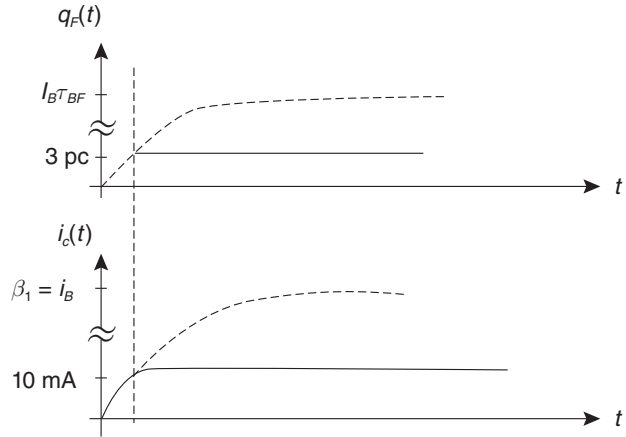
$$\frac{dq_F}{dt} + \frac{q_F}{\tau_{BF}} = I_B u_{-1}(t) \quad [10-23]$$

where $\tau_{BF} = \beta_F \tau_F = 125$ nanoseconds and $u_{-1}(t)$ is the unit step. This charge control equation has the solution:

$$q_F(t) = I_B \tau_{BF} \left(1 - e^{-\frac{t}{\tau_{BF}}} \right) \quad [10-24]$$

$$i_c(t) \approx \frac{q_F(t)}{\tau_F} \approx \beta_F I_B \left(1 - e^{-\frac{t}{\tau_{BF}}} \right) = (416 \text{ mA}) \left(1 - e^{-\frac{t}{125 \text{ ns}}} \right)$$

Figure 10-11: Growth of forward charge $q_F(t)$ and collector current $i_c(t)$ when crossing the forward-active region. The final value of collector current is 10 milliamps, at which point the transistor enters saturation and the forward charge q_F and collector current remain constant.



Note that the collector current rises with characteristic time constant $\tau_{BF} = 125$ nanoseconds. The final value of collector current is approximately 10 milliamps (at which point the transistor saturates). This final value of collector current is reached at $t_{ri} \approx 3$ nanoseconds and at this time the forward charge $q_F = i_c \tau_F = 3$ picocoulombs. Of course, the actual time to cross the forward active region will be somewhat higher due to base current needed to charge the base-emitter and base-collector junction capacitances.

Going into saturation

Since $\beta_F I_B > I_{c(sat)}$, the transistor does not remain forward active for all time. Let's find the point at which saturation occurs. Then we'll find the final values (in saturation) of the saturation charge⁸ q_S and evaluate the time constant τ_S . The charge control equations for base and collector current in saturation are given by:⁹

$$i_b - \frac{I_{c,SAT}}{\beta_F} = \frac{q_S}{\tau_S} + \frac{dq_S}{dt'} \quad [10-25]$$

$$i_c = I_{c,SAT}$$

⁸ We'll assume that the base current drive remains on long enough to fully saturate the transistor, and to fully fill up the final value of saturation charge.

⁹ Note that we've changed the time scale; we assume now that we enter saturation at $t' = 0$. We'll change the time scale throughout the problem as we progress.

The time constant in the saturation region is:

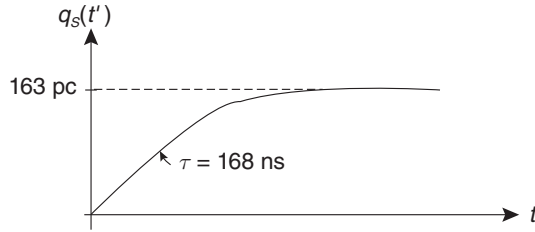
$$\tau_s = \frac{\tau_{BF}(\beta_R + 1) + \tau_{BR}(\beta_F + 1)}{\beta_F + \beta_R + 1} = \frac{(125 \text{ ns})(0.7 + 1) + (168 \text{ ns})(416 + 1)}{416 + 0.7 + 1} = 168 \text{ ns} \quad [10-26]$$

The final value of saturation charge q_{sfinal} is found by:

$$q_{sfinal} = \left(I_B - \frac{I_{C,SAT}}{\beta_F} \right) \tau_s = \left(1 \text{ mA} - \frac{10 \text{ mA}}{416} \right) (168 \text{ ns}) = 163 \text{ pC} \quad [10-27]$$

This saturation charge (163 picocoulombs) is the charge that needs to be removed from the base before the transistor will leave saturation, when we try to turn it off.

Figure 10-12: Growth of saturation charge $q_s(t')$ while going into saturation. During this interval the collector current is constant at 10 milliamps.



Leaving saturation

Now, after waiting a long time, the base drive is turned off, i.e., $i_B(t) = 0$. Let's figure out how long the transistor remains saturated, with $i_C = I_{c(sat)}$. Since the transistor is saturated, but the base drive value is zero, the charge control equation is now:¹⁰

$$-\frac{I_{C,SAT}}{\beta_F} = \frac{q_s}{\tau_s} + \frac{dq_s}{dt''} \quad [10-28]$$

Note that recombination in the base causes the saturation charge to decrease. This equation needs to be solved with initial value of $q_s(t'')$ being 163 picocoulombs. We need to solve this equation and find out how long it takes $q_s(t'')$ to drop to zero.¹¹ A solution of this equation is:

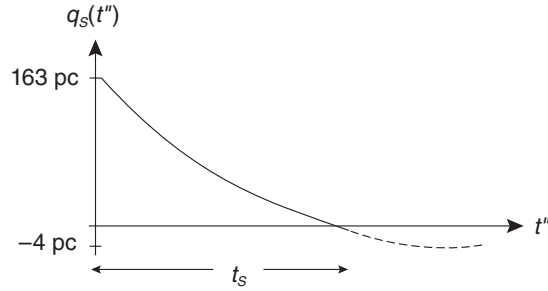
$$q_s(t'') = 163 \text{ pC} - (167 \text{ pC}) \left(1 - e^{\frac{-t''}{\tau_s}} \right) \quad [10-29]$$

This equation shows that $q_s(t'')$ drops to zero at $t'' = 627$ nanoseconds, which is when the transistor leaves saturation. Therefore, our saturation delay time is $t_s = 627$ nanoseconds (**Figure 10-13**).

¹⁰ We've again changed the time scale.

¹¹ If we let the charge control equation run forever, the equation says that the final value of q_s will be -4 nanocoulombs. However, we know that the transistor leaves saturation when the saturation charge drops to zero. We'll find the time when q_s drops to zero.

Figure 10-13: Decay of saturation charge $q_s(t'')$ while coming out of saturation. During this interval, the collector current is constant at 10 milliamps.



Crossing through the forward active region again while turning off

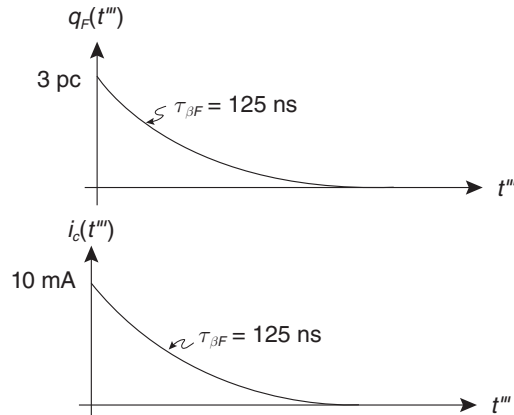
The transistor enters the forward active region again when q_s is zero. We'll next find $q_F(t''')$ and the collector current. We expect the turn-off time of the collector current to be significantly slower than the turn-on time, since we aren't actively pulling any base current out of the base. We have to rely on recombination in the base to decrease $q_F(t''')$ and hence collector current. The charge control equation in this case is:

$$\frac{dq_F}{dt'''} + \frac{q_F}{\tau_{BF}} = 0 \quad [10-30]$$

Note that the right-hand side of Eq. 10-30 is zero, since there is no base current. At the beginning of the forward active region, the collector current is 10 milliamps and the forward charge $q_F = I_c \tau_F = (10 \text{ mA})(0.3 \text{ ns}) = 3 \text{ picocoulombs}$. The solution to this charge control equation is:

$$\begin{aligned} q_F(t) &= q_{Fo} e^{\frac{-t'''}{\tau_{BF}}} = (3 \text{ pC}) e^{\frac{-t'''}{125 \text{ ns}}} \\ i_c(t) &= \frac{q_F(t''')}{\tau_F} = (10 \text{ mA}) e^{\frac{-t'''}{125 \text{ ns}}} \end{aligned} \quad [10-31]$$

Figure 10-14: Decay of forward charge $q_F(t''')$ and resultant collector current $i_c(t''')$ when crossing the forward-active region and turning the transistor off. The collector current decays to zero with characteristic time constant $\tau_{\beta F}$.



Therefore, once the transistor leaves saturation the collector current will decay with characteristic time constant of 125 nanoseconds. The SPICE plot below shows a saturation turn-off delay time of approximately 500 nanoseconds, and a slowly decaying collector current. This PSPICE simulation is reasonably consistent with what we calculated.

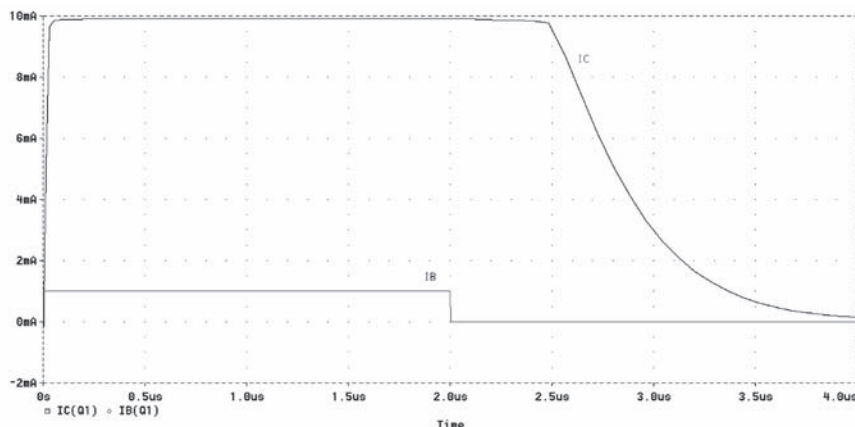


Figure 10-15: Transistor with base current drive (SPICE solution).
In this simulation the current pulse is turned on at $t = 0$, and turned off at $t = 2 \mu\text{s}$.

Example 10.2: Transistor inverter with voltage drive

Consider the transistor inverter shown in **Figure 10-16a**. Assume that we use a transistor with the following specifications:

- $f_T = 300 \text{ MHz}$
- $\beta_F = 200$
- $\tau_R = 10 \text{ ns}$
- $\beta_R = 10$
- For junction capacitances, use the 2N3904 junction capacitances found in **Figure 4-25**.

We will use the charge control model to find the approximate switching profile.

There are six regions of operation:

- 1) Traversing the cutoff region (t_{dl}). During this time, the nonlinear depletion region capacitances C_{je} and C_{jc} are charged to the value that puts the transistor on the beginning of the forward active region. We can approximate the voltage when the transistor begins to carry collector current as $V_{BE} \approx 0.4\text{V}$.
- 2) Traversing the forward active region (t_r). During this time, V_{BE} increases from 0.4V to approximately 0.7V, and the collector current rises to its final value.
- 3) Going into hard saturation. When $V_{CB} \approx 0$, the transistor enters the saturation region. In hard saturation, excess base charge builds up in the base as the base-collector junction becomes more and more forward-biased.

After we have waited a long time so that the transistor is fully in saturation, we have the turn-off waveforms:

- 4) Turn-off (saturation) delay (t_{sd}) when saturation charge is removed. In this period, the collector current remains approximately constant.
- 5) Traversing the forward-active region (t_f) where the collector current drops back to zero.
- 6) Traversing the cutoff region. The space charge layer capacitances are discharged as V_{BE} moves to its final value. During this time, the collector current is approximately zero.

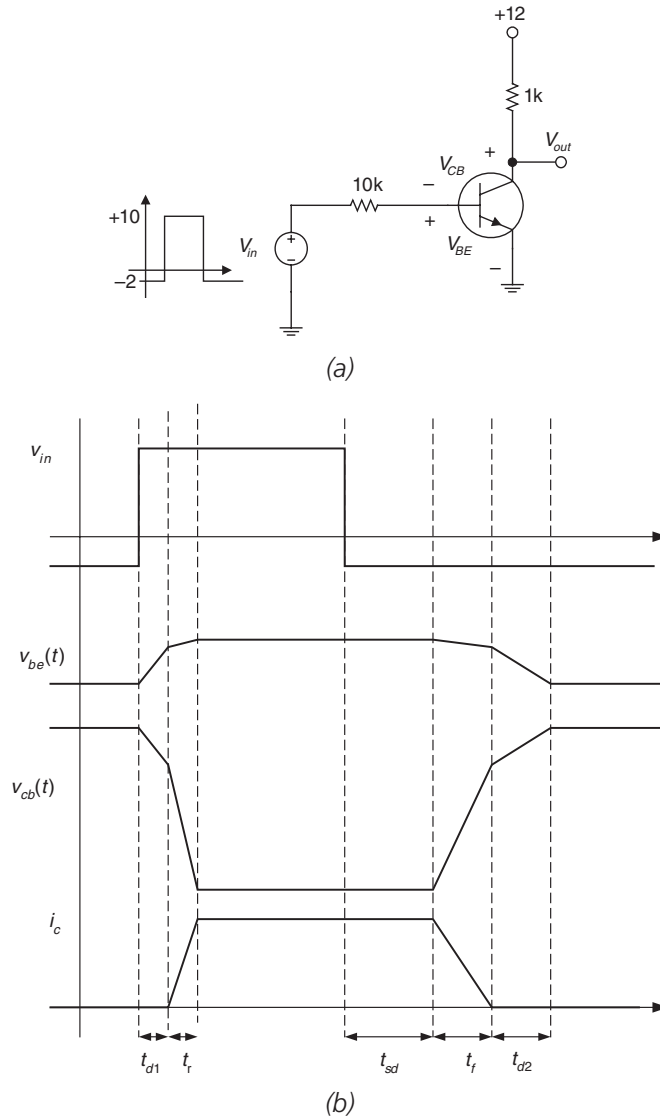


Figure 10-16: Charge control example. (a) Circuit (b) Switching waveforms.

Calculations for the various regions are as follows:

Cutoff region: At $t = 0$, we assume that the input instantaneously transitions from -2V to $+10\text{V}$. The turn-on delay time t_{d1} is the time that it takes to charge the nonlinear collector-base and base-emitter capacitances so that $V_{BE} = 0.4\text{V}$, which we will take as the beginning of the forward-active region. During this charging interval, the base current does vary a little bit, since the V_{BE} changes, but for simplicity we'll use the average base current to determine the switching time. Terminal conditions are:

At the beginning of the switching interval, at $t = 0$:

$$V_{BE} = -2\text{V}$$

$$V_{CB} = +14$$

$$C_{je} = 2.9 \text{ picofarads @ } V_{BE} = -2\text{V}$$

$$q_{ve} = C_{je} V_{BE} = (2.9 \text{ picofarads})(-2\text{V}) = -5.8 \text{ picocoulombs}$$

$$C_{jc} = 1.5 \text{ picofarads @ } V_{CB} = +14\text{V}$$

$$q_{vc} = C_{jc} V_{CE} = (1.5 \text{ picofarads})(14\text{V}) = 21 \text{ picocoulombs}$$

$$i_b(t = 0) = (10\text{V} - -2\text{V})/10\text{k} = 1.2 \text{ mA}$$

$$i_c(t = 0) = 0$$

At the very beginning of the forward active region, $t = t_{d1}$:

$$V_{BE} = 0.4\text{V}$$

$$V_{CB} = +11.6$$

$$C_{je} \approx 8 \text{ picofarads @ } V_{BE} = 0.4 \text{ V; (this is an estimate, since the curves on the datasheet don't extend to positive voltages).}$$

$$q_{ve} = C_{je} V_{BE} = (8 \text{ picofarads})(0.4\text{V}) = +3.2 \text{ picocoulombs}$$

$$C_{jc} = 1.6 \text{ picofarads @ } V_{CB} = +11.6\text{V}$$

$$q_{vc} = C_{jc} V_{CE} = (1.6 \text{ picofarads})(11.6\text{V}) = 18.6 \text{ picocoulombs}$$

$$i_b(t = t_{d1}) = (10\text{V} - 0.4\text{V})/10\text{k} = 0.96 \text{ mA}$$

$$i_c(t = t_{d1}) = 0$$

The average base current $i_{b,avg} = 1.08$ milliamps during this charging interval in the cutoff region. The change in charge in the nonlinear base-emitter capacitance is $\Delta q_{ve} = (5.8 \text{ picocoulombs} + 3.2 \text{ picocoulombs})$, resulting in $\Delta q_{ve} = 9 \text{ picocoulombs}$. A similar calculation for the collector-base junction results in $\Delta q_{vc} = 2.4 \text{ picocoulombs}$. The time to charge is approximated by:

$$t_{d1} \approx \frac{\Delta q_{ve} + \Delta q_{vc}}{i_{b,avg}} = \frac{10.4 \text{ pC}}{1.08 \text{ mA}} \approx 10 \text{ ns} \quad [10-32]$$

During this interval, the collector current is zero.

Traversing the forward active region: At $t = t_{d1}$, we assume that the transistor begins to turn on and the collector current begins to rise, until saturation is reached. The rise time t_r depends on the rate that we can both charge the nonlinear junction capacitances, as well as supply the

excess base charge that supports the collector current. During the forward active region, the charge control equations are:

$$\begin{aligned} i_c &\approx \frac{q_F}{\tau_F} \\ i_b &= \frac{q_F}{\tau_{BF}} + \frac{d}{dt}(q_F + q_{ve} + q_{vc}) \end{aligned} \quad [10-33]$$

At the beginning of saturation, we'll assume that $V_{CB} = 0$ and $V_{BE} = 0.7$, requiring $i_{c,sat} = 11.3$ milliamps. Remember that the forward transit time is related to the transistor f_T by $\tau_F = 1/(2\pi f_T) = 0.53$ ns. Therefore, the forward charge required is $q_F = i_{c,sat}\tau_F = (11.3 \text{ mA})(0.53 \text{ ns}) = 6$ picocoulombs. This is the base charge needed to completely turn the transistor on, until it begins to saturate. An additional charge is needed to charge C_{je} from 0.4V to 0.7V and C_{jc} from 11.6V to 0V. The calculations are as follows:

At $t = t_{d1}$:

$$\begin{aligned} V_{BE} &= 0.4\text{V} \\ V_{CB} &= 11.6 \\ C_{je} &\approx 8 \text{ picofarads @ } V_{BE} = 0.4 \text{ V} \\ C_{jc} &= 1.6 \text{ picofarads @ } V_{CB} = 11.6\text{V} \\ i_b(t = t_{d1}) &= 0.96 \text{ mA} \end{aligned}$$

At $t = t_{d1} + t_r$:

$$\begin{aligned} V_{BE} &= 0.7\text{V} \\ V_{CB} &= 0 \\ C_{je} &\approx 8 \text{ picofarads @ } V_{BE} = 0.7 \text{ V} \\ C_{jc} &= 3.5 \text{ picofarads @ } V_{CB} \approx 0 \\ i_b(t = t_{d1}) &= 0.93 \text{ mA} \end{aligned}$$

The average base current $i_{b,avg} = 0.945$ mA during this charging interval, and the change in charge in the capacitors is $\Delta q_{ve} = 2.4$ picocoulombs and $\Delta q_{vc} = 18.6$ picocoulombs. Therefore, the switching time is (approximately):

$$t_r \approx \frac{q_F + \Delta q_{ve} + \Delta q_{vc}}{i_{b,avg}} = \frac{27 \text{ pC}}{0.945 \text{ mA}} = 28.6 \text{ ns} \quad [10-34]$$

Going into saturation: At $t = t_{d1} + t_{r1}$, the transistor enters saturation and the transistor V_{BE} is clamped at approximately 0.7V. The charge control equations in saturation are:

$$\begin{aligned} i_c &= I_{C,sat} \\ i_b - \frac{I_{C,sat}}{\beta_F} &= \frac{dq_s}{dt} + \frac{q_s}{\tau_s} \end{aligned} \quad [10-35]$$

The second term in the second charge control equation for base current ($I_{C,sat}/\beta_F$) is the base current needed to support recombination in the base. In the saturation region, the collector current is constant, and the base saturation charge builds up with a time constant given by:

$$\tau_s = \frac{\tau_{BF}(\beta_R + 1) + \tau_{BR}\beta_F}{\beta_F + \beta_R + 1} \quad [10-36]$$

The total saturation charge once the transistor is in full saturation is given by:

$$q_s = \left(i_b - \frac{I_{C,sat}}{\beta_F} \right) \tau_s \quad [10-37]$$

For this example, $\tau_s = 100$ nanoseconds and $q_s = 87.6$ picocoulombs. The q_s is the saturation charge that must be removed before the transistor can be brought out of saturation.

Saturation storage (turn-off) delay: During this interval, the collector current remains constant at $I_{C,sat}$ while the saturation charge is removed. The saturation charge q_s is removed by two mechanisms: reverse base current and by recombination. The charge-control equations are the same as in the previous case:

$$\begin{aligned} i_c &= I_{C,sat} \\ i_b - \frac{I_{C,sat}}{\beta_F} &= \frac{dq_s}{dt} + \frac{q_s}{\tau_s} \end{aligned} \quad [10-38]$$

During the saturation turn-off delay, the input voltage is $-2V$ and the V_{BE} is clamped at approximately $0.7V$; hence the base current $i_b = -0.27$ mA. The recombination current is $I_{C,sat}/\beta_F = 11.3$ mA/200 = 0.056 mA. The base current charge control equation is:

$$-0.326 \text{ mA} = \frac{dq_s}{dt} + \frac{q_s}{\tau_s} \quad [10-39]$$

Solving this charge control equation for the time when q_s drops to zero results in $t_{sd} = 131$ nanoseconds, as shown in **Figure 10-17**.

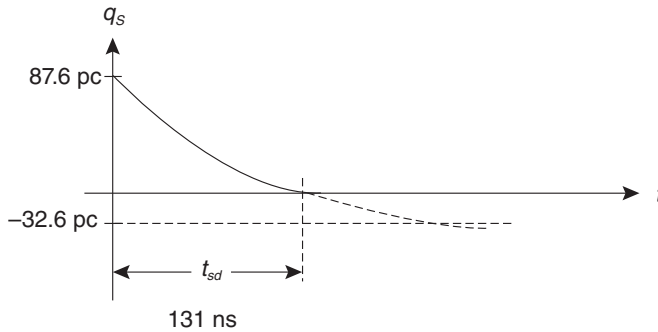


Figure 10-17: Decay of saturation charge q_s while coming out of saturation. During this interval, the collector current is constant and the saturation charge decays from 87.6 picocoulombs to zero.

Forward-active region (current falltime): After leaving saturation, we again enter the forward-active region. In order to turn the transistor off and traverse the forward-active region, we need to remove the base charge q_F as well as discharge the nonlinear depletion capacitances. At the beginning of the current falltime interval:

$$V_{BE} = 0.7\text{V}$$

$$V_{CB} \approx 0$$

$$C_{je} \approx 8 \text{ picofarads @ } V_{BE} = 0.7\text{V}$$

$$C_{jc} = 3.5 \text{ picofarads @ } V_{CB} = 0$$

$$i_b = 0.27 \text{ mA}$$

At the end of the falltime interval:

$$V_{BE} = 0.4\text{V}$$

$$V_{CB} = 11.6$$

$$C_{je} \approx 8 \text{ picofarads @ } V_{BE} = 0.4\text{V}$$

$$C_{jc} = 1.6 \text{ picofarads @ } V_{CB} = 11.6\text{V}$$

$$i_b = 0.24 \text{ mA}$$

The base charge q_F that we need to remove is the same as in the turn-on case ($q_F = 6$ picocoulombs). The change in charge in the depletion regions is also the same as in the turn-on case, resulting in:

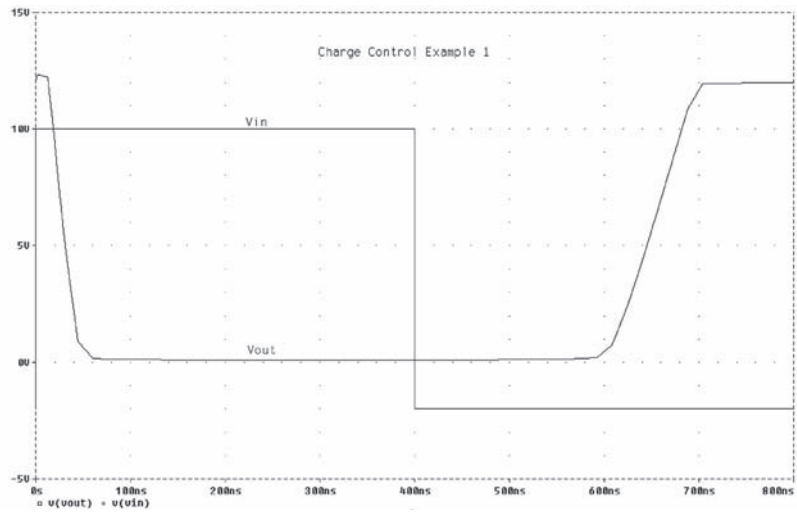
$$t_f \approx \frac{q_F + \Delta q_{ve} + \Delta q_{vc}}{i_{b,avg}} = \frac{27 \text{ pC}}{0.255 \text{ mA}} = 106 \text{ ns} \quad [10-40]$$

Note that even though the total charge supplied is the same as in the turn-on case, the turn-off time is much longer since the average base current is lower.

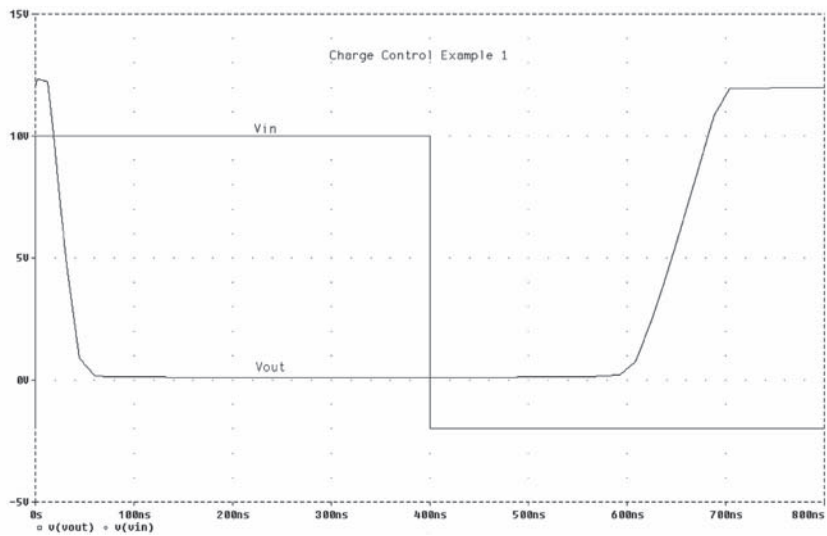
Cutoff: During this interval, we discharge the nonlinear capacitances back to their final values. The total gate charge required is the same as the turn-on case, where the charge is 11.44 picocoulombs. In this case the average base current over the interval is approximately 0.12 mA, resulting in a total turnoff delay time of:

$$t_{d2} \approx \frac{\Delta q_{ve} + \Delta q_{vc}}{i_{b,avg}} = \frac{11.44 \text{ pC}}{0.12 \text{ mA}} = 95.3 \text{ ns} \quad [10-41]$$

A SPICE simulation for the inverter is shown in **Figure 10-18**. **Table 10-1** shows a comparison of the charge control calculations with the SPICE calculations. This result shows that the charge control model is a useful approximation technique for calculating the switching times. More importantly, the model shows how to speed up the switching process. In this example, the output risetime and falltime can be significantly reduced by reducing the base resistor.



(a)



(b)

Figure 10-18: SPICE result for charge control example (a) Input and output voltages. Note that the saturation delay time is roughly 200 ns. (b) Base current.

Table 10-1: Comparison of Charge Control Model Results with SPICE.

Item	Charge control	SPICE	% diff
1. Cutoff	1.06E-08	1.25E-08	15.3%
2. FAR	2.85E-08	3.15E-08	9.4%
3. Saturation			
4. Storage	1.31E-07	1.90E-07	31.3%
5. FAR	1.06E-07	1.00E-07	-5.7%
6. Cutoff	9.53E-08		

Now, in order to speed up the switching, we reduce the base resistor from 10k to 1k (**Figure 10-19**). As expected, the risetime and falltime of the output is much faster. However, the saturation delay time is approximately the same as the previous example. Why?

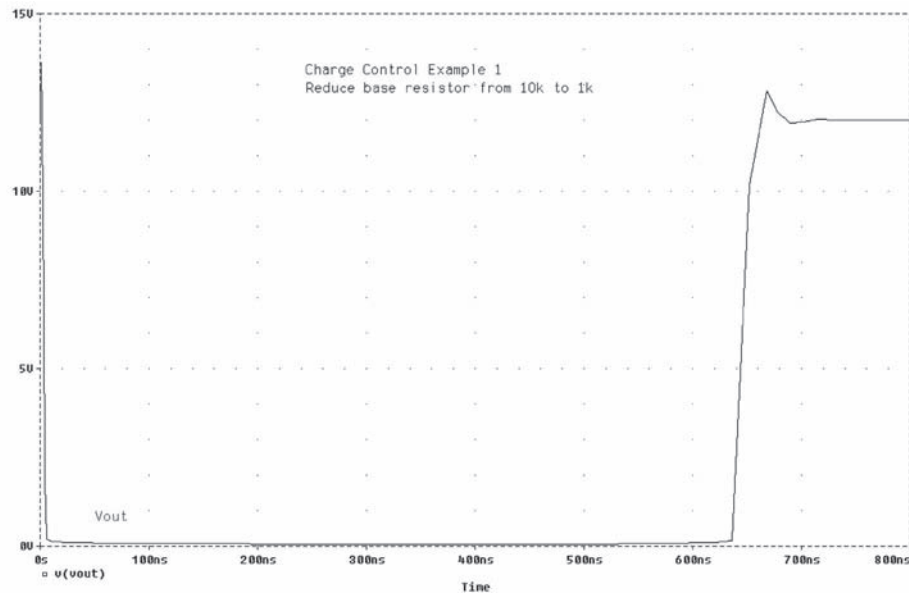


Figure 10-19: Output waveform when base resistor is reduced from 10k to 1k. The risetime and falltime are much faster. Note that the saturation delay time is still roughly 200 ns.

We can figure out why the saturation delay time remains approximately constant by considering the charge control equation for the transistor in saturation. Note that the final value of saturation charge q_s depends on how hard we drive the base current of the transistor. This results in a saturation delay time that is approximately constant.

Example 10.3: Nonsaturating current switch

A nonsaturating current switch implemented with 2N2222 signal transistors is shown in **Figure 10-20**. If we ignore junction capacitances, the relevant charge control equations for the base and collector currents are:

$$i_b = \frac{q_F}{\tau_{BF}} + \frac{dq_F}{dt} \quad [10-42]$$

$$i_c = \frac{q_F}{\tau_F}$$

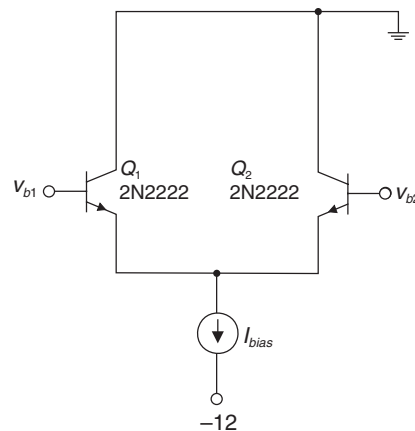
In order to turn the transistors ON and OFF fast, we need to supply base current to charge and discharge q_F for each transistor. Therefore, in large part, the switching speed is determined by how much base current we can supply. The base current in this voltage-driven case is set by the voltage swing of the base drivers and the base spreading resistance r_x of the transistors. For the 2N2222 transistor,¹² the value of the base spreading resistance is very roughly $r_x \approx 40\Omega$.

We'll assume a 2V swing for the base drivers $v_{b1}(t)$ and $v_{b2}(t)$, so the maximum base current is ~50 mA. We can solve for the switching time as:

$$\tau_{sw} \approx \frac{q_{F,final}}{\langle i_b \rangle} \approx \frac{I_{C,final} \tau_F}{\langle i_b \rangle} \approx \frac{I_{C,final}}{\omega_T \langle i_b \rangle} \quad [10-43]$$

where the notation $\langle i_b \rangle$ denotes the average value of base current during switching. The actual switching time will be longer, due to the fact that the same base current must charge the nonlinear base-collector and base-emitter junction capacitances. The important result from this is that the switching time will approximately scale with the collector current, if the base current remains constant. **Figure 10-21** is a transient SPICE simulation for I_{BIAS} levels of 100 mA, 200 mA and 300 mA. We note that the switching time is a few nanoseconds¹³ in each

Figure 10-20: Differential current switch.



¹² This is a crude approximation since the value of the base spreading resistance varies with collector current bias level.

¹³ Note that a nanosecond is a very short period of time. Light travels about a foot in a nanosecond.

case and this switching time does increase approximately proportionally with the bias current level, as expected. **Figure 10-22** is a transient SPICE plot of base current. During the transition of the collector current, the base currents are high.

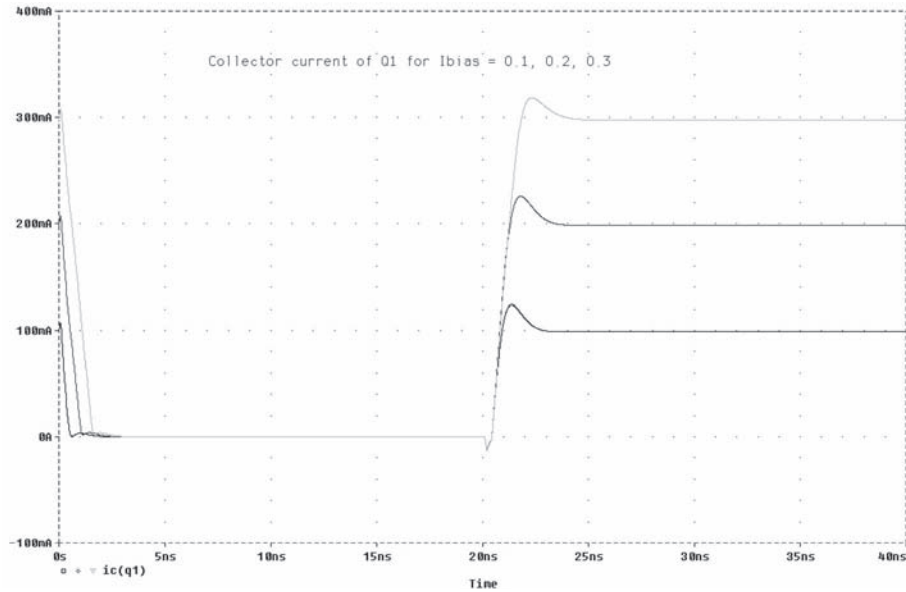


Figure 10-21: Current switch simulation, showing collector current of Q_1 .

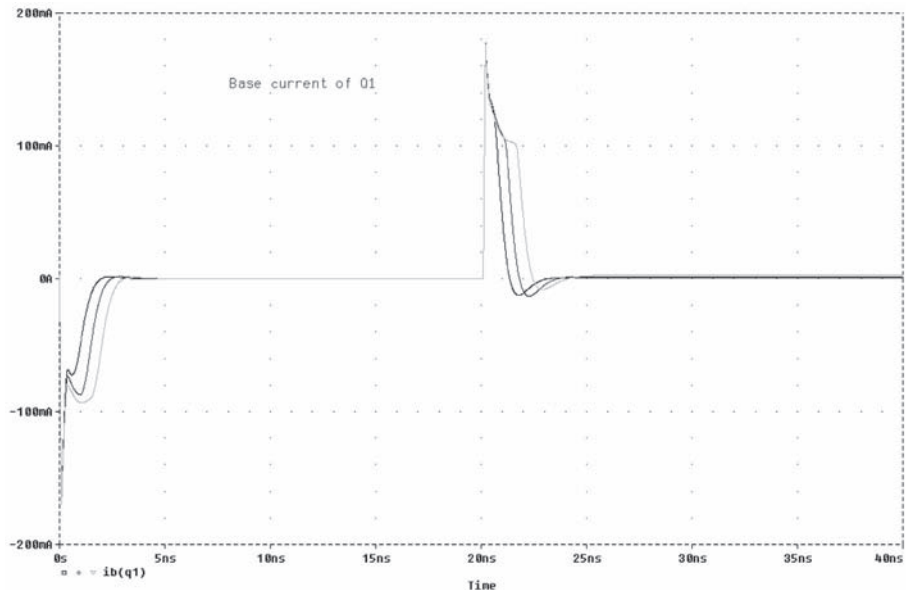


Figure 10-22: Current switch simulation, showing base current of Q_1 .

Emitter Switching

An emitter switch is shown in **Figure 10-23**. This type of topology has been used in power switches where high voltage isolation to the load is needed. As shown in **Figure 10-23b**, the MOSFET provides the high speed current switching, and the bipolar transistor is a common-base buffer which provides voltage isolation. Note that in the emitter switching case, there is no problem with transistor saturation. Hence, we expect this type of switch to be fast. We'll estimate the switching speed of this type of topology assuming that the MOSFET provides a step of current to the emitter of the transistor.

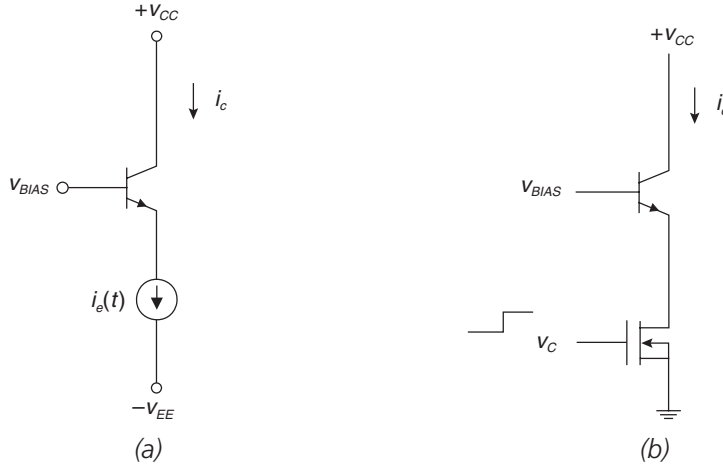


Figure 10-23: Emitter switching. (a) Circuit. (b) Implementation as a high voltage switch with the low-side MOSFET providing the fast switching current, and the transistor providing the high-voltage buffering.

Referring to **Figure 10-23a** and ignoring junction capacitances, the relevant charge control equations are:

$$\begin{aligned}
 i_b &= \frac{q_F}{\tau_{BF}} + \frac{dq_F}{dt} \\
 i_c &= \frac{q_F}{\tau_F} \\
 i_e &= -(i_b + i_c) = -\frac{q_F}{\tau_{BF}} - \frac{dq_F}{dt} - \frac{q_F}{\tau_F}
 \end{aligned} \tag{10-44}$$

Rearranging the above emitter charge control equation results in:

$$-\frac{q_F}{\tau_{BF}} - \frac{dq_F}{dt} - \frac{q_F}{\tau_F} = -I_E u_{-1}(t) \tag{10-45}$$

where $u_{-1}(t)$ is the unit step. We can make a further approximation for the emitter charge control equation by noting that $\tau_{BF} \gg \tau_F$ and hence $q_F/\tau_{BF} \ll q_F/\tau_F$, resulting in:

$$\frac{dq_F}{dt} + \frac{q_F}{\tau_F} \approx I_E u_{-1}(t) \quad [10-46]$$

The solution to this charge control equation is a simple exponential rise in transistor forward charge q_F with characteristic time constant τ_F .

$$q_F(t) \approx I_E \tau_F \left(1 - e^{-\frac{t}{\tau_F}} \right) \quad [10-47]$$

$$i_c = \frac{q_F(t)}{\tau_F} \approx I_E \left(1 - e^{-\frac{t}{\tau_F}} \right)$$

This solution shows that the collector current rises with characteristic time τ_F , which is comparable to $1/\omega_T$ of the transistor, as shown in the SPICE plot in **Figure 10-24**.

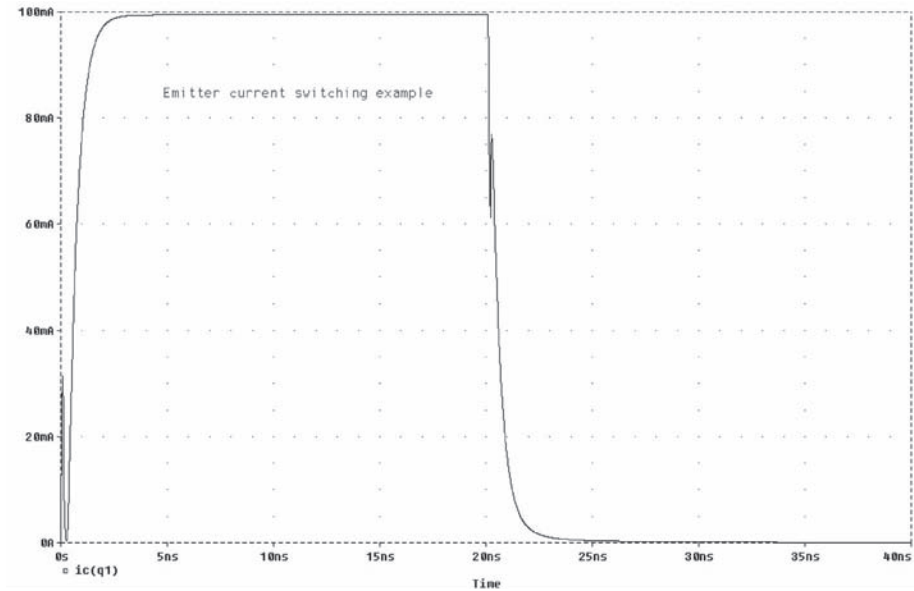


Figure 10-24: Emitter switch—PSPICE simulation result, using 2N3904 transistor.

2N2222 Datasheet Excerpts¹⁴

Amplifier Transistors
NPN Silicon

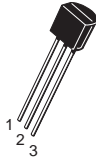
MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Collector–Emitter Voltage	V_{CEO}	40	Vdc
Collector–Base Voltage	V_{CBO}	75	Vdc
Emitter–Base Voltage	V_{EBO}	6.0	Vdc
Collector Current — Continuous	I_C	600	mAdc
Total Device Dissipation @ $T_A = 25^\circ\text{C}$ Derate above 25°C	P_D	625 5.0	mW mW/ $^\circ\text{C}$
Total Device Dissipation @ $T_C = 25^\circ\text{C}$ Derate above 25°C	P_D	1.5 12	Watts mW/ $^\circ\text{C}$
Operating and Storage Junction Temperature Range	T_J, T_{stg}	-55 to $+150$	$^\circ\text{C}$

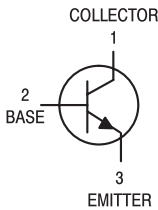
THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction to Ambient	$R_{\theta JA}$	200	$^\circ\text{C}/\text{W}$
Thermal Resistance, Junction to Case	$R_{\theta JC}$	83.3	$^\circ\text{C}/\text{W}$

P2N2222A



CASE 29–11, STYLE 17
TO–92 (TO–226AA)



ELECTRICAL CHARACTERISTICS ($T_A = 25^\circ\text{C}$ unless otherwise noted)

Characteristic	Symbol	Min	Max	Unit
OFF CHARACTERISTICS				
Collector–Emitter Breakdown Voltage ($I_C = 10$ mAdc, $I_B = 0$)	$V_{(BR)CEO}$	40	—	Vdc
Collector–Base Breakdown Voltage ($I_C = 10$ μAdc , $I_E = 0$)	$V_{(BR)CBO}$	75	—	Vdc
Emitter–Base Breakdown Voltage ($I_E = 10$ μAdc , $I_C = 0$)	$V_{(BR)EBO}$	6.0	—	Vdc
Collector Cutoff Current ($V_{CE} = 60$ Vdc, $V_{EB(off)} = 3.0$ Vdc)	I_{CEX}	—	10	nAdc
Collector Cutoff Current ($V_{CB} = 60$ Vdc, $I_E = 0$) ($V_{CB} = 60$ Vdc, $I_E = 0$, $T_A = 150^\circ\text{C}$)	I_{CBO}	— —	0.01 10	μAdc
Emitter Cutoff Current ($V_{EB} = 3.0$ Vdc, $I_C = 0$)	I_{EBO}	—	10	nAdc
Collector Cutoff Current ($V_{CE} = 10$ V)	I_{CEO}	—	10	nAdc
Base Cutoff Current ($V_{CE} = 60$ Vdc, $V_{EB(off)} = 3.0$ Vdc)	I_{BEX}	—	20	nAdc

¹⁴ Reprinted with permission of On Semiconductor.

ELECTRICAL CHARACTERISTICS ($T_A = 25^\circ\text{C}$ unless otherwise noted) (Continued)

Characteristic	Symbol	Min	Max	Unit
ON CHARACTERISTICS				
DC Current Gain ($I_C = 0.1\text{ mA}$, $V_{CE} = 10\text{ Vdc}$) ($I_C = 1.0\text{ mA}$, $V_{CE} = 10\text{ Vdc}$) ($I_C = 10\text{ mA}$, $V_{CE} = 10\text{ Vdc}$) ($I_C = 10\text{ mA}$, $V_{CE} = 10\text{ Vdc}$, $T_A = -55^\circ\text{C}$) ($I_C = 150\text{ mA}$, $V_{CE} = 10\text{ Vdc}$) ⁽¹⁾ ($I_C = 150\text{ mA}$, $V_{CE} = 1.0\text{ Vdc}$) ⁽¹⁾ ($I_C = 500\text{ mA}$, $V_{CE} = 10\text{ Vdc}$) ⁽¹⁾	h_{FE}	35 50 75 35 100 50 40	— — — — 300 — —	—
Collector–Emitter Saturation Voltage ⁽¹⁾ ($I_C = 150\text{ mA}$, $I_B = 15\text{ mA}$) ($I_C = 500\text{ mA}$, $I_B = 50\text{ mA}$)	$V_{CE(sat)}$	— —	0.3 1.0	Vdc
Base–Emitter Saturation Voltage ⁽¹⁾ ($I_C = 150\text{ mA}$, $I_B = 15\text{ mA}$) ($I_C = 500\text{ mA}$, $I_B = 50\text{ mA}$)	$V_{BE(sat)}$	0.6 —	1.2 2.0	Vdc

SMALL-SIGNAL CHARACTERISTICS

Current–Gain — Bandwidth Product ⁽²⁾ ($I_C = 20\text{ mA}$, $V_{CE} = 20\text{ Vdc}$, $f = 100\text{ MHz}$)	f_T	300	—	MHz
Output Capacitance ($V_{CB} = 10\text{ Vdc}$, $I_E = 0$, $f = 1.0\text{ MHz}$)	C_{obo}	—	8.0	pF
Input Capacitance ($V_{EB} = 0.5\text{ Vdc}$, $I_C = 0$, $f = 1.0\text{ MHz}$)	C_{ibo}	—	25	pF
Input Impedance ($I_C = 1.0\text{ mA}$, $V_{CE} = 10\text{ Vdc}$, $f = 1.0\text{ kHz}$) ($I_C = 10\text{ mA}$, $V_{CE} = 10\text{ Vdc}$, $f = 1.0\text{ kHz}$)	h_{ie}	2.0 0.25	8.0 1.25	k Ω
Voltage Feedback Ratio ($I_C = 1.0\text{ mA}$, $V_{CE} = 10\text{ Vdc}$, $f = 1.0\text{ kHz}$) ($I_C = 10\text{ mA}$, $V_{CE} = 10\text{ Vdc}$, $f = 1.0\text{ kHz}$)	h_{re}	— —	8.0 4.0	$\times 10^{-4}$
Small–Signal Current Gain ($I_C = 1.0\text{ mA}$, $V_{CE} = 10\text{ Vdc}$, $f = 1.0\text{ kHz}$) ($I_C = 10\text{ mA}$, $V_{CE} = 10\text{ Vdc}$, $f = 1.0\text{ kHz}$)	h_{fe}	50 75	300 375	—
Output Admittance ($I_C = 1.0\text{ mA}$, $V_{CE} = 10\text{ Vdc}$, $f = 1.0\text{ kHz}$) ($I_C = 10\text{ mA}$, $V_{CE} = 10\text{ Vdc}$, $f = 1.0\text{ kHz}$)	h_{oe}	5.0 25	35 200	μmhos
Collector Base Time Constant ($I_E = 20\text{ mA}$, $V_{CB} = 20\text{ Vdc}$, $f = 31.8\text{ MHz}$)	r_b/C_c	—	150	ps
Noise Figure ($I_C = 100\text{ }\mu\text{A}$, $V_{CE} = 10\text{ Vdc}$, $R_S = 1.0\text{ k}\Omega$, $f = 1.0\text{ kHz}$)	N_F	—	4.0	dB

SWITCHING CHARACTERISTICS

Delay Time	(V _{CC} = 30 Vdc, V _{BE(off)} = –2.0 Vdc, I _C = 150 mA, I _{B1} = 15 mA) (Figure 1)	t _d	—	10	ns
Rise Time		t _r	—	25	ns
Storage Time	(V _{CC} = 30 Vdc, I _C = 150 mA, I _{B1} = I _{B2} = 15 mA) (Figure 2)	t _s	—	225	ns
Fall Time		t _f	—	60	ns

1. Pulse Test: Pulse Width $\leq 300\text{ }\mu\text{s}$, Duty Cycle $\leq 2.0\%$.

2. f_T is defined as the frequency at which $|h_{fe}|$ extrapolates to unity.

P2N2222A

SWITCHING TIME EQUIVALENT TEST CIRCUITS

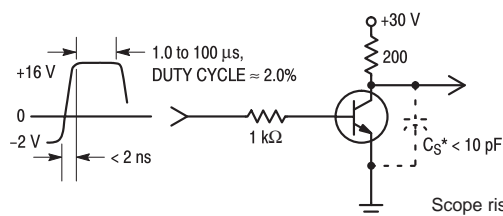


Figure 1. Turn-On Time

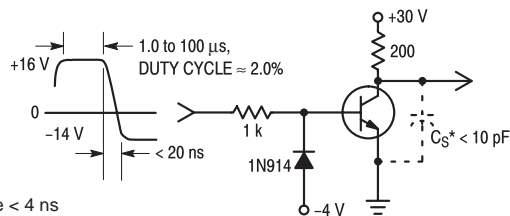


Figure 2. Turn-Off Time

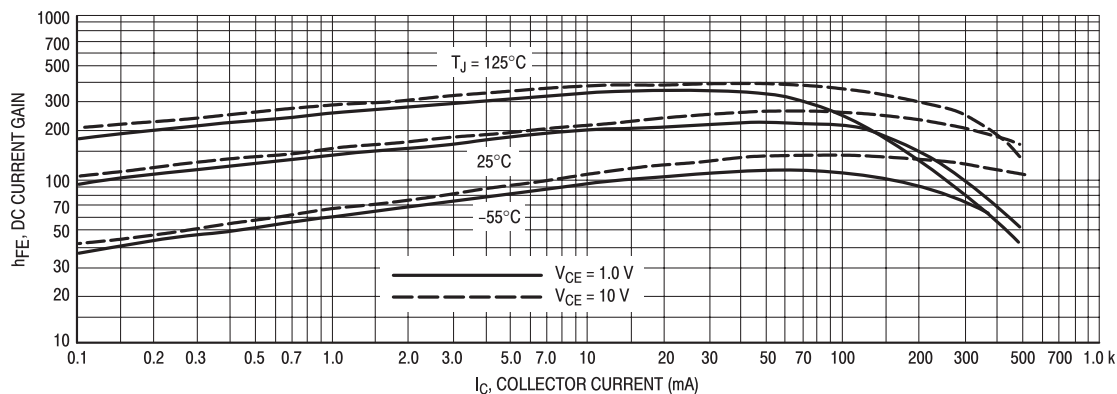


Figure 3. DC Current Gain

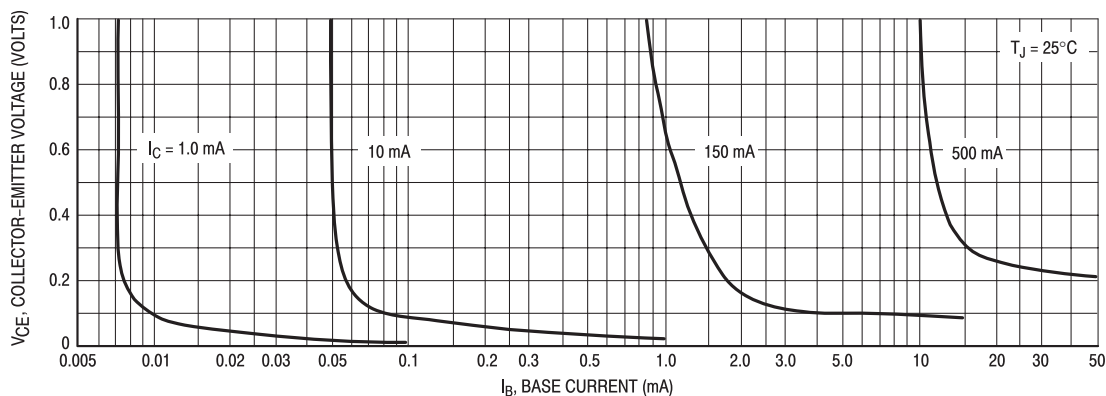


Figure 4. Collector Saturation Region

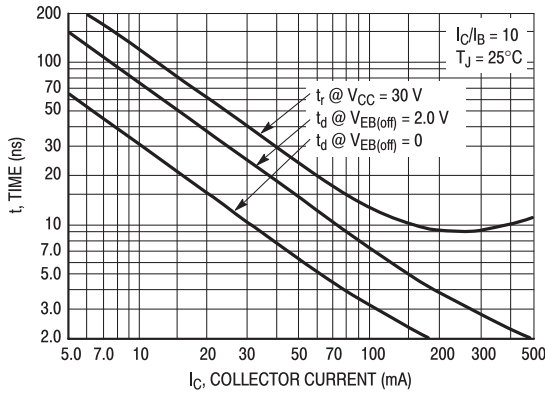


Figure 5. Turn-On Time

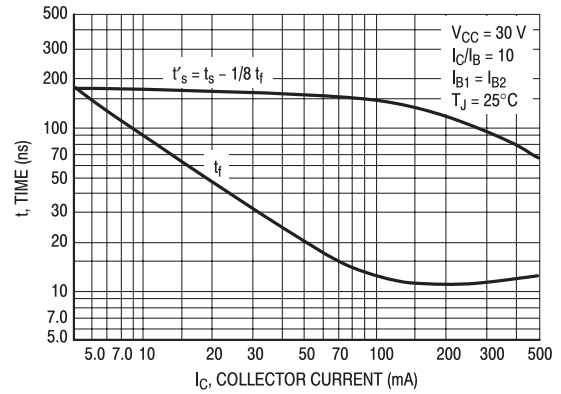


Figure 6. Turn-Off Time

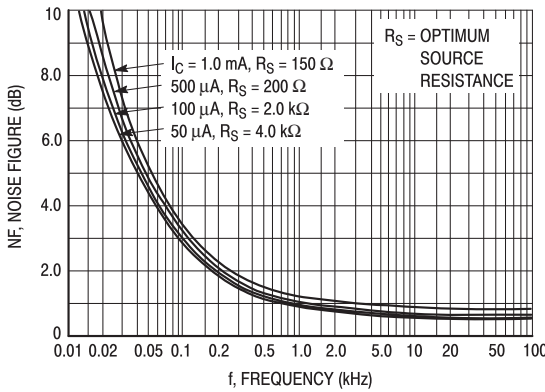


Figure 7. Frequency Effects

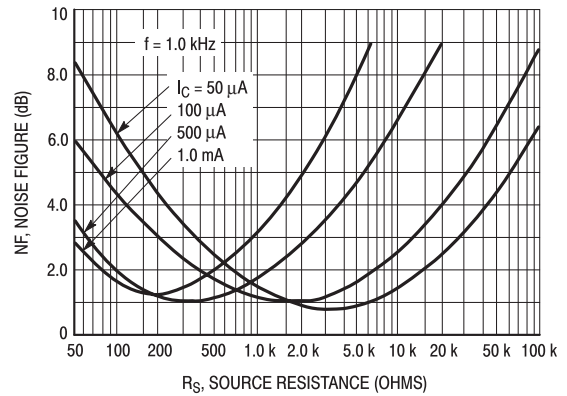


Figure 8. Source Resistance Effects

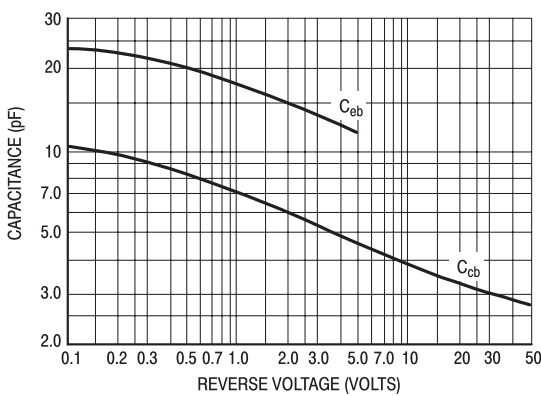


Figure 9. Capacitances

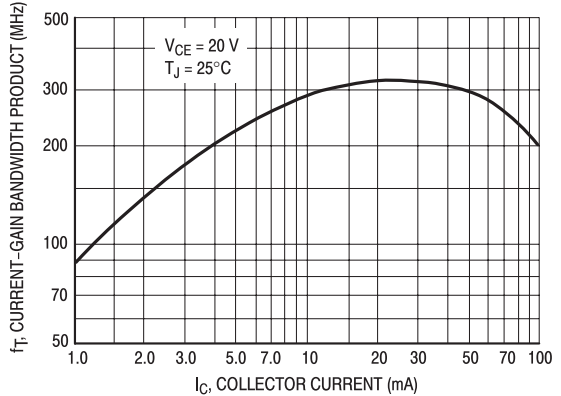


Figure 10. Current-Gain Bandwidth Product

Chapter 10 Problems

Problem 10.1

A pulsed transistor circuit is shown in **Figure 10-25a**. The transistor is driven by a pulse $v_i(t)$ (**Figure 10-25b**) which transitions from -5V to $+5\text{V}$ at time $t = 0$. At time $t = 10\text{ }\mu\text{s}$, the input pulse $v_i(t)$ transitions from -5V to $+5\text{V}$. Assume that when the transistor is saturated that $V_{BE} = 0.7\text{V}$. Transistor parameters are as follows: $\tau_F = 0.5\text{ ns}$; $\beta_F = 50$; $\tau_R = 50\text{ ns}$; $\beta_R = 5$. The junction capacitances can be found from the following:

$$C_j = \frac{C_{jo}}{\sqrt{1 - \frac{V_j}{\phi_{bi}}}}$$

For each junction $\phi_{bi} = 0.9\text{V}$. For the base-emitter junction $C_{jo} = 35\text{ pF}$. For the base-collector junction, $C_{jo} = 10\text{ pF}$. Plots of C_{je} and C_{jc} are shown in **Figure 10-26**.

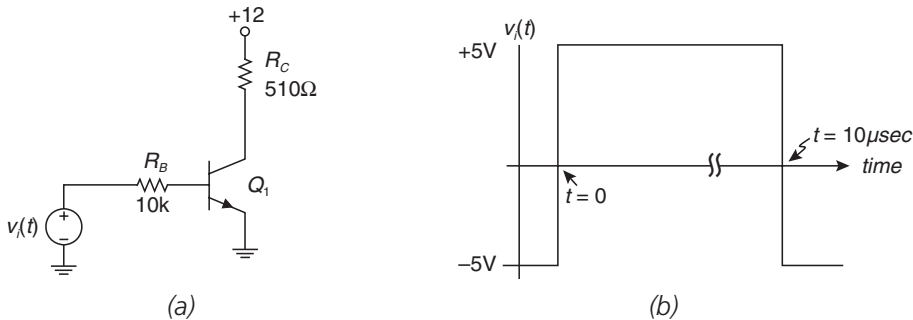


Figure 10-25: Transistor circuit for Problem 10.1. (a) Circuit. (b) Base drive pulse.

- At $t = 0+$, the input pulse $v_i(t)$ transitions high. Approximately how long does the transistor remain in cutoff before entering the forward active region? In doing this part of the problem, identify all terminal conditions that help you determine the switching time.
- After leaving the cutoff region (near $V_{BE} = 0.7\text{V}$) the transistor enters the forward-active region. Determine the time to cross the forward-active region $t_{far,ON}$. Sketch $v_{be}(t)$ and collector current $i_c(t)$ in the interval $0 < t < 500\text{ ns}$, indicating cutoff time and collector current risetime

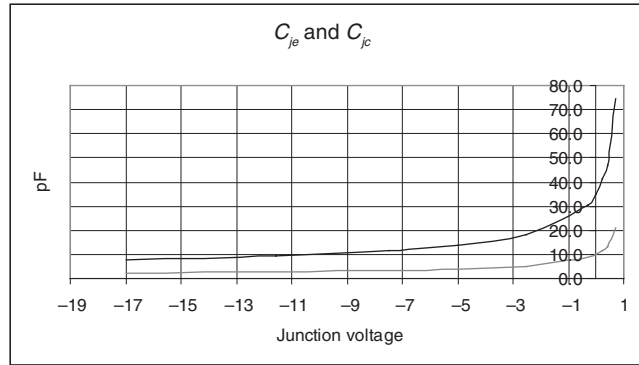


Figure 10-26: Junction capacitances for Problem 10.1.

- (c) At time $t = 10 \mu\text{s}$, the input pulse transitions from +5V to -5V. Find and sketch the collector current for $t > 10 \mu\text{s}$.

Problem 10.2

The amplifier in **Figure 10-27** is driven by a 10.6V voltage step at $t = 0$.

- (a) Using the charge-control model, sketch $i_c(t)$ and $v_{out}(t)$ for the transistor amplifier shown below. Include the effect of space-charge capacitances. Make reasonable assumptions, including using average values for base current, capacitances, etc. For the 2N3904, assume that $\beta_F = 100$.
- (b) After you finish the hand calculations using the charge control model, simulate your circuit using SPICE. Compare your results and explain any discrepancies. Note: for the input voltage, use a SPICE pulse generator with a risetime much faster ($>10\times$) than the dynamics you expect. Make the duration of the transient simulation run long enough so that you see the final value of collector current and output voltage.

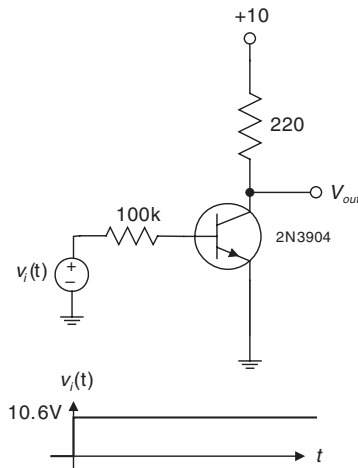


Figure 10-27: Transistor switch for Problem 10.2.

Problem 10.3

In this problem, a 2N3904 transistor is controlled by a voltage drive as shown in **Figure 10-28**. Analyze the dynamics of the transistor. Assume transistor parameters: $\beta_F = 100$. Ignore space-charge capacitances.

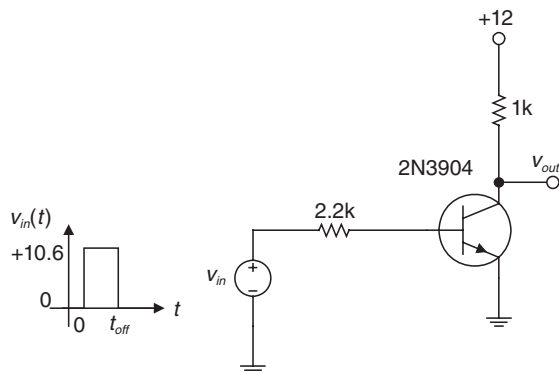


Figure 10-28: Transistor switch for Problem 10.3.

- Assume that the transistor remains in the forward active region for all time. Determine the time constants and final values, etc., and sketch $q_F(t)$.
- How long (t_d) does it take to traverse the active region on turn ON?
- Since $\beta_F i_B > I_{C(sat)}$, the transistor will not remain forward active for all time. Indicate on your graphs the point at which saturation occurs. Find q_{BO} , the value of q_F that puts the transistor on the edge of saturation. Find the value of $I_{C,SAT}$ evaluate the time constant τ_S .
- Now, after waiting a long time, the transistor is fully saturated. After reaching full saturation, what is the value of saturation charge?
- After the transistor has reached full saturation, the input pulse goes low. Obviously, the transistor does not remain saturated, but enters the forward active region when q_S is zero. Determine the storage delay time, the time during which the device remains saturated after the input voltage pulse drops to zero. Sketch the saturation charge q_S .
- Compare turn-on and turn-off times, and explain the difference.
- Simulate your circuit using PSPICE and compare your charge control result with the simulation results. Attempt to explain any major discrepancies.
- It is noted that the storage delay time decreases if the input pulse duration is reduced. Explain.
- Now, the transistor is driven with voltage pulse shown in **Figure 10-29**.

Simulate your circuit using PSPICE and explain any major differences in the switching waveforms as compared to part (g).

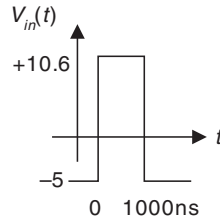


Figure 10-29: New drive waveform for Problem 10.3.

Problem 10.4

A transistor is in a switching circuit as shown in **Figure 10-30**. At time $t = 0$ the base drive current source transitions instantaneously from 0 mA to 10 mA. The current source remains energized for 1000 ns, after which it turns OFF. The transistor has $\omega_T = 10^9$ radians/sec, $\beta_F = 100$; $\beta_R = 5$; $\tau_R = 50$ ns; $V_{CE,SAT} = 0$. Assume that the transistor enters the forward-active region at $V_{BE} = 0.5$ V. Junction capacitance values are shown in the chart below.

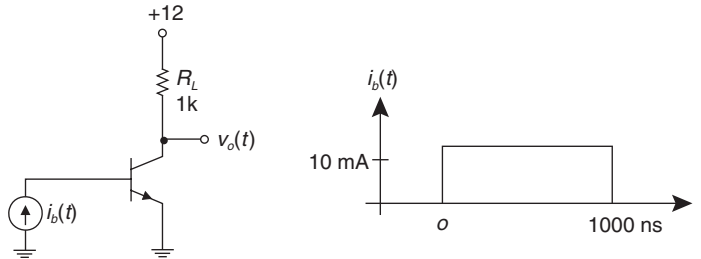


Figure 10-30: Transistor switch for Problem 10.4.

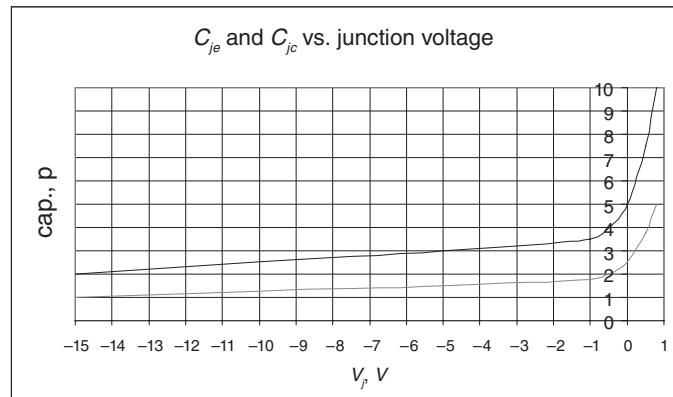


Figure 10-31: Junction capacitances vs. junction voltage. Top trace: C_{je} . Bottom trace: C_{jc} .

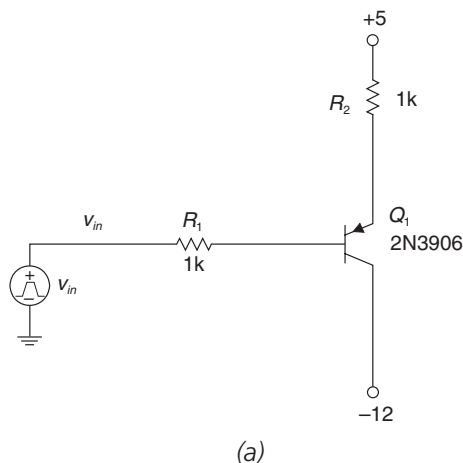
- At $t = 0+$, the current source transitions HIGH. find the transistor turn-ON delay t_d .
- After crossing the cutoff region, the transistor enters the forward-active region. Calculate the approximate risetime of collector current t_{ri} .

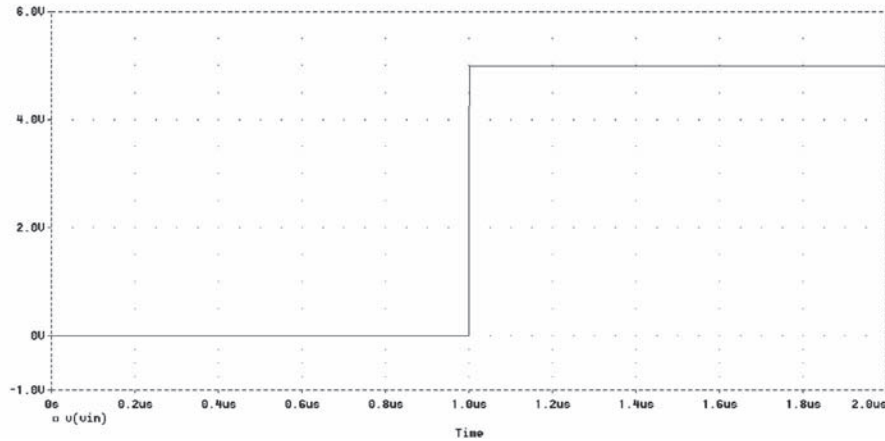
- (c) After crossing the forward-active region, the transistor enters saturation. Calculate the final value of saturation charge q_s .
- (d) After the current source turns OFF, the transistor remains saturated for a time. Calculate the storage delay time t_s .
- (e) After the transistor leaves saturation, it again crosses the forward active region when the collector current drops to zero. Calculate the approximate current falltime t_f . (In this section, estimate the falltime as the time it takes the collector current to fall from 100% of initial value to 10% of initial value.)
- (f) Sketch the collector current and transistor V_{CE} , labeling all axes.

Problem 10.5

For this transistor in **Figure 10-32**, use: $C_{je} = 10$ pF, $C_{jc} = 2$ pF, $f_T = 300$ MHz, $\beta_F = 100$, $\beta_R = 5$, $\tau_R = 50$ ns.

- (a) The circuit in **Figure 10-30** is driven by a 0V to 5V step at 1000 ns. What is the initial collector current at $t = 0$?
- (b) Does the transistor ever saturate, and why or why not?
- (c) At $t = 1000$ ns, the input voltage steps from 0V to 5V. Calculate the delay time before the collector current begins changing and also calculate the transition time of collector current approximately as the transistor transitions through the forward active region. State your assumptions and back them up.
- (d) Sketch the input waveform and the transistor collector current.





(b)

Figure 10-32: Circuit for Problem 10.5.

References

- Barna, A., "Analytic approximations for propagation delays in current-mode switching circuits including collector-base capacitances," *IEEE Journal of Solid-State Circuits*, vol. 16, no. 5, October 1981, pp. 597–599.
- Bashkow, Theodore R., "Effect of Nonlinear Collector Capacitance on Collector Current Rise Time," *IRE Transactions on Electron Devices*, October 1956, pp. 167–172.
- Casaravilla, G., and Silveira, F., "Emitter drive: a technique to drive a bipolar power transistor switching at 100 kHz," *Proceedings of the 1990 IEEE Colloquium in South America*, August 31–September 15, 1990, pp. 188–192.
- Chuang, C. T., and Chin, K., "High-speed low-power direct-coupled complementary push-pull ECL circuit," *IEEE Journal of Solid-State Circuits*, vol. 29, no. 7, July 1994, pp. 836–839.
- Easley, James W., "The Effect of Collector Capacity on the Transient Response of Junction Transistors," *IRE Transactions on Electron Devices*, January 1957, pp. 6–14.
- Ebers, J. J., and Moll, J. L., "Large-Signal Behavior of Junction Transistors," *Proceedings of the I.R.E.*, 1954, p. 1761.
- , "Large-Signal Behavior of Junction Transistors," *Proceedings of the IRE*, December 1954, pp. 1761–1772.
- Ghannam, M. Y., Mertens, R. P., and van Overstraeten, R. J., "An analytical model for the determination of the transient response of CML and ECL gates," *IEEE Transactions on Electron Devices*, vol. 37, no. 1, January 1990, pp. 191–201.
- Gray, Paul E., and Searle, Campbell L., *Electronic Principles Physics, Models and Circuits*, John Wiley, 1969.

- Jensen, Randall W., "Charge Control Transistor Model for the IBM Electronic Circuit Analysis Program," *IEEE Transactions on Circuit Theory*, vol. CT-13, no. 4, December 1966, pp. 428–437.
- Karadzinov, L. V., Arsov, G. L., Dzekov, T. A., and Jeffries, D. J., "Charge-control piecewise-linear bipolar junction transistor model," *Proceedings of the IEEE International Symposium on Industrial Electronics*, ISIE '96, June 17–20, 1996, pp. 561–566.
- Konstadinidis, G. K., and Berger, H. H., "Optimization of buffer stages in bipolar VLSI systems," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 7, July 1992, pp. 1002–1013.
- Kuno, H. John, "Rise and Fall Time Calculations of Junction Transistors," *IEEE Transactions on Electron Devices*, April 1964, pp. 151–155.
- Moll, John L., "Large-Signal Transient Response of Junction Transistors," *Proceedings of the I.R.E.*, December 1954, pp. 1773–1784.
- Muller, Richard S., and Kamins, Theodore I., *Device Electronics for Integrated Circuits*, 2d edition, John Wiley, 1986.
- Musumeci, S., Pagano, R., Raciti, A., Porto, C., Ronsisvalle, C., and Scollo, R., "Characterization, parameter identification and modeling of a new monolithic emitter-switching bipolar transistor," *Proceedings of the 39th Industry Applications Conference*, October 3–7, 2004, pp. 1924–1931.
- Neudeck, Gerold W., "The PN Junction Diode," *Modular Series on Solid State Devices*, volume II, Addison-Wesley, 1983.
- Robinson, F. V. P., and Williams, B. W., "Optimising snubbers for high-current emitter-switched transistors," *Proceedings of the Third International Conference on Power Electronics and Variable-Speed Drives*, July 13–15, 1988, pp. 177–180.
- Sharaf, K. M., and Elmasry, M. I., "An accurate analytical propagation delay model for high-speed CML bipolar circuits," *IEEE Journal of Solid-State Circuits*, vol. 29, no. 1, January 1994, pp. 31–45.
- Stork, J. M. C., "Bipolar transistor scaling for minimum switching delay and energy dissipation," *Technical Digest of International Electron Devices Meeting*, December 11–14, 1988, pp. 550–553.
- Thompson, M. T., and Schlecht, M. F., "High power laser diode driver based on power converter technology," *IEEE Transactions on Power Electronics*, vol. 12, no. 1, January 1997, pp. 46–52.
- Winkel, Jan TE, "Extended Charge-Control Model for Bipolar Transistors," *IEEE Transactions on Electron Devices*, vol. ED-20, no. 4, April 1973, pp. 389–394.

Review of Feedback Systems

In This Chapter

- The arena of “classical control system” covers single-input, single-output (SISO) linear time invariant (LTI), continuous time systems. This includes the design of classical servomechanisms and most operational amplifier circuits. This chapter offers introductory material for the design of feedback control systems.

Introduction and Some Early History of Feedback Control

A feedback system is one that compares its output to a desired input and takes corrective action to force the output to follow the input. Arguably, the beginnings of automatic feedback control¹ can be traced back to the work of James Watt in the 1700s. Watt did lots of work on steam engines, and he adapted² a centrifugal *governor* to automatically control the speed of a steam engine. The governor was comprised of two rotating metal balls that would fly out due to centrifugal force. The amount of “fly-out” was then used to regulate the speed of the steam engine by adjusting a throttle. This was an example of proportional control.

The steam engines of Watt’s day worked well with the governor, but as steam engines became larger and better engineered, it was found that there could be stability problems in the engine speed. One of the problems was *hunting*, or an engine speed that would surge and decrease, apparently hunting for a stable operating point. This phenomenon was not well understood until the latter part of the 19th century, when James Maxwell³ (yes, the same Maxwell famous for all those equations) developed the mathematics of the stability of the Watt governor using differential equations.

¹ Others may argue that the origins of feedback control trace back to the water clocks and float regulators of the ancients. See, e.g., Otto Mayr’s *The Origins of Feedback Control*, The MIT Press, 1970.

² The centrifugal governor was invented by Thomas Mead c. 1787, for which he received British Patent #1628.

³ James C. Maxwell, “On Governors,” *Proceedings of the Royal Society*, 1867, pp. 270–283.

Invention of the Negative Feedback Amplifier

We now jump forward to the 20th century. In the early days of telephone, practical difficulties were encountered in building a transcontinental telephone line. The first transcontinental telephone system, built in 1914, used #8 copper wire weighing about 1000 pounds per mile. Loss due to the resistance of the wire was approximately 60dB.⁴ Several vacuum tube amplifiers were used to boost the amplitude. These amplifiers have limited bandwidth and significant nonlinear distortion. The effects of cascading amplifiers (**Figure 11-1**) resulted in intolerable amounts of signal distortion.

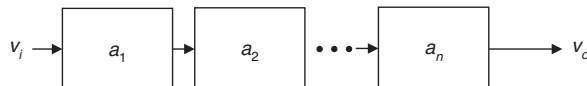


Figure 11-1: Amplifier cascade

Harold Black graduated from Worcester Polytechnic Institute in 1921 and joined Bell Laboratory. At this time, a major task facing AT&T was the improvement of the telephone system and the problem of distortion in cascaded amplifiers. In 1927, Black⁵ was considering the problem of distortion in amplifiers and came up with the idea of the negative feedback amplifier.

“Then came the morning of Tuesday, August 2, 1927, when the concept of the negative feedback amplifier came to me in a flash while I was crossing the Hudson River on the Lackawanna Ferry, on the way to work. For more than 50 years I have pondered how and why the idea came, and I can’t say any more today than I could that morning. All I know is that after several years of hard work on the problem, I suddenly realized that if I fed the amplifier output back to the input, in reverse phase, and kept the device from oscillating (singing, as we called it then), I would have exactly what I wanted: a means of canceling out the distortion in the output. I opened my morning newspaper and on a page of The New York Times I sketched a simple diagram of a negative feedback amplifier plus the equations for the amplification with feedback. I signed the sketch, and 20 minutes later, when I reached the laboratory at 463 West Street, it was witnessed, understood, and signed by the late Earl C. Blessing.

I envisioned this circuit as leading to extremely linear amplifiers (40 to 50dB of negative feedback), but an important question is: How did I know I could avoid self-oscillations over very wide frequency bands when many people doubted such circuits would be stable? My confidence stemmed from work that I had done two years earlier on certain novel oscillator circuits and three years earlier in designing the terminal circuits, including the filters, and developing the mathematics for a carrier telephone system for short toll circuits.”

⁴ William McC. Siebert, *Circuits, Signals and Systems*, The MIT Press, 1986.

⁵ Harold Black, “Inventing the Negative Feedback Amplifier,” *IEEE Spectrum*, December 1977, pp. 55–60. See also Harold Black’s U.S. Patent #2,102, 671, “Wave Translation System,” filed April 22, 1932 and issued December 21, 1937, and Black’s early paper “Stabilized Feed-Back Amplifiers,” *Bell System Technical Journal*, 1934.

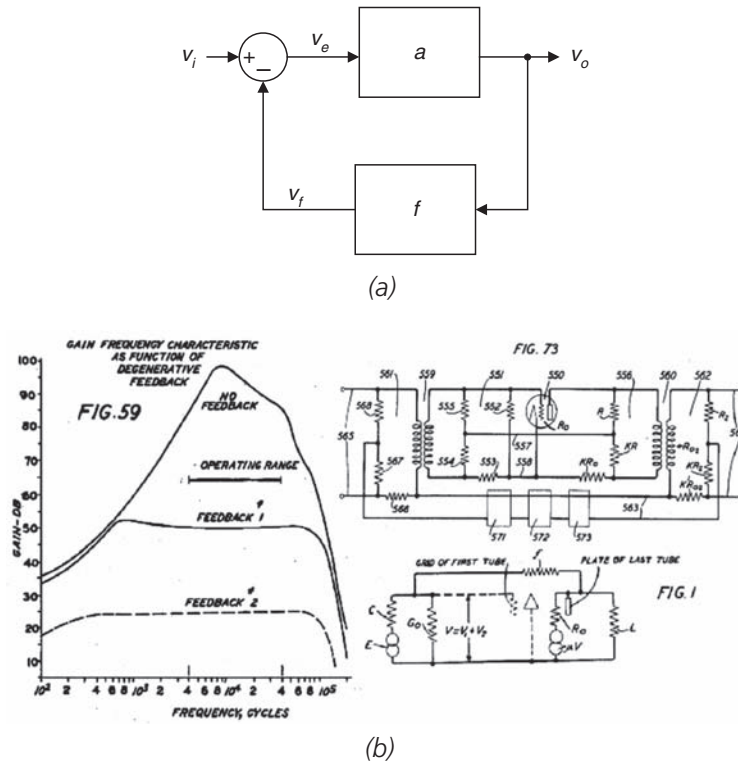
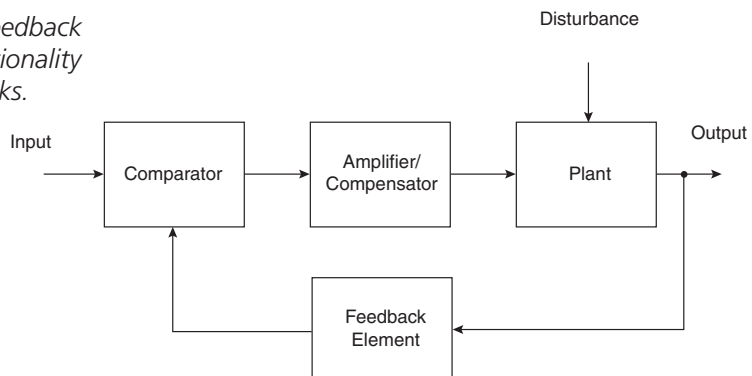


Figure 11-2: Classical single-input, single-output control loop, as envisioned by Black.
 (a) Block diagram form. (b) Excerpt from Black's U.S. patent #2,102,671 issued in 1937.

A typical closed-loop negative feedback system as is commonly implemented is shown in **Figure 11-3**. The “plant” in this diagram might represent, for instance, the power stage in an audio amplifier. A properly designed control system can maintain the output at a desired level in the face of external disturbances and uncertainties in the model of the plant. The goal of the feedback system is to force the output to track the input, perhaps with some gain and frequency-response shaping.

Figure 11-3: Typical feedback system showing functionality of individual blocks.



In this configuration, the output signal is fed back to the input, where it is compared with the desired input. The difference between the two signals is amplified and applied to the plant input.

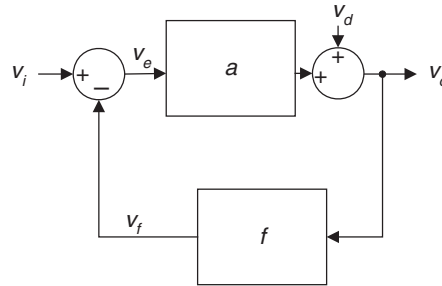
In order to design a successful feedback system, several issues must be resolved.

- First, how do you generate the model of the plant, given that many systems do not have well-defined transfer functions?
- Once you have the model of the plant, how do you close the loop, resulting in a stable system with a desired gain and bandwidth?

Control System Basics

A classical feedback loop, as envisioned by Black, is shown in **Figure 11-4**. Note that there is an external disturbance in this system, the voltage v_d .

Figure 11-4: Classical single input, single output control loop, with input voltage v_i , output voltage v_o and external disturbance v_d .



In this system, a is the forward path gain and f is the feedback gain. The forward gain a and feedback factor f may have frequency dependence (and, hence, the plant should be denoted as $a(s)$) but for notational simplicity we'll drop the Laplace variable s .

Initially, let's set the disturbance v_d to zero. The “error” term v_e is the difference between the input and the fed-back portion of the output. We can solve for the transfer function with the result:

$$\begin{aligned} v_o &= av_e \\ v_e &= v_i - v_f \\ v_f &= fv_o \end{aligned} \tag{11-1}$$

The closed-loop gain is: ($A \equiv$ closed-loop gain):

$$A = \frac{v_o}{v_i} = \frac{a}{1 + af} \tag{11-2}$$

or:

$$A = \frac{\text{FORWARD GAIN}}{1 - \text{LOOP TRANSMISSION}} \tag{11-3}$$

Note what happens in the limit of $af \gg 1$:

$$A \approx \frac{1}{f} \tag{11-4}$$

This is the key to designing a successful feedback system; if you can guarantee that $af \gg 1$ for the frequencies that you are interested in, *then your closed-loop gain will not be dependent on the details of the plant gain $a(s)$* . This is very useful, since in some cases the feedback function f can be implemented with a simple resistive divider which can be cheap and accurate.

Loop Transmission and Disturbance Rejection

The term in the denominator of the gain equation is $1 + af$, where the term $-af$ is called the loop transmission. This term is the gain going around the whole feedback loop; you can find the L.T. by doing a thought experiment: Cut the feedback loop in one place, inject a signal, and find out what returns where you cut. The gain around the loop is the loop transmission.

Now, let's find the gain from the disturbance input to the output:

$$\frac{v_o}{v_d} = \frac{1}{1 + af} = \frac{1}{1 - L.T.} \quad [11-5]$$

Note that if the loop transmission is large at frequencies of interest, then the output due to the disturbance will be small. The term $(1 + af)$ is called the *desensitivity* of the system. Let's figure out the fractional change in closed-loop gain A due to a change in forward path gain a .

$$\begin{aligned} A &= \frac{a}{1 + af} \\ \frac{dA}{da} &= \frac{(1 + af) - af}{(1 + af)^2} = \left(\frac{1}{1 + af} \right) \left(\frac{1}{1 + af} \right) = \frac{A}{a} \left(\frac{1}{1 + af} \right) \\ \therefore \frac{dA}{A} &= \frac{da}{a} \left(\frac{1}{1 + af} \right) \end{aligned} \quad [11-6]$$

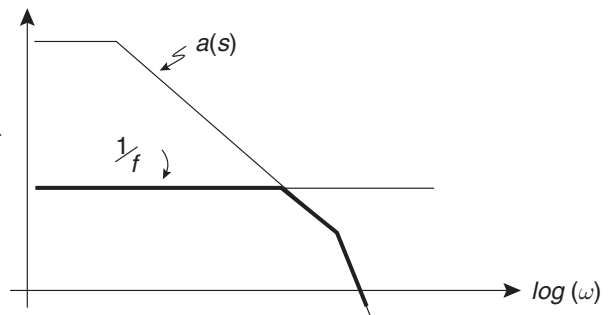
This result means that if $af \gg 1$ then the fractional change in closed-loop gain (dA/A) is much smaller than the fractional change in forward-path gain (da/a).

We can make a couple of approximations in the limit of large and small loop transmission.

For large loop transmission ($af \gg 1$), as we've shown before, the closed loop gain $A \approx 1/f$.

For small loop transmission ($af \ll 1$), the closed-loop gain is approximately $a(s)$. If we plot $a(s)$ and $1/f$ on the same set of axes, we can find an approximation for the closed loop gain as the lower of the two curves, as shown in **Figure 11-5**.

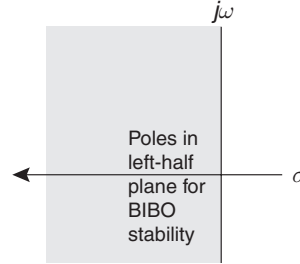
Figure 11-5: Plot for estimating closed-loop transfer function graphically. The curve $a(s)$ depicts the frequency dependence of the forward-path gain. The line $1/f$ is the inverse of the feedback gain, shown here for resistive feedback. The thick line indicates our estimate for closed-loop transfer function. For $a(s)f \gg 1$, the closed-loop gain is approximately $1/f$. For $a(s)f \ll 1$, the closed-loop gain is approximately $a(s)$.



Stability

So far, we haven't discussed the issue regarding the stability of closed-loop systems. There are many definitions of stability in the literature, but we'll consider "BIBO" stability. In other words, we'll consider the stability problem given that we'll only excite our system with bounded inputs. The system is BIBO stable if *bounded inputs* generate *bounded outputs*, a condition that is met if all poles are in the left-half plane (**Figure 11-6**).

Figure 11-6: Closed-loop pole locations in the left-half plane for bounded input, bounded output (BIBO) stability.

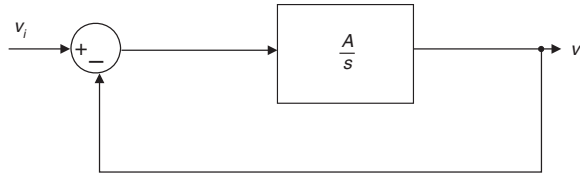


Consider the feedback system with a first-order plant and unity feedback (**Figure 11-7**). The input/output transfer function is:

$$\frac{v_o}{v_i} = \frac{\frac{A}{s}}{\frac{A}{s} + 1} = \frac{A}{s + A} = \frac{1}{\frac{s}{A} + 1} \quad [11-7]$$

Note that as the forward-path gain A increases, the closed-loop bandwidth increases as well, with the closed-loop pole staying on the real axis at $s = -A$. As long as A is positive, this system is BIBO stable for any values of A .

Figure 11-7: First-order system comprised of an integrator inside a negative-feedback loop.



The second-order system (**Figure 11-8a**) is also easy to work out, with transfer function:

$$\frac{v_o}{v_i} = \frac{\frac{K}{(\tau_a s + 1)(\tau_b s + 1)}}{1 + \frac{K}{(\tau_a s + 1)(\tau_b s + 1)}} = \frac{K}{K + (\tau_a s + 1)(\tau_b s + 1)} = \left(\frac{K}{K + 1} \right) \frac{1}{\left(\frac{\tau_a \tau_b}{1 + K} \right) s^2 + \left(\frac{\tau_a + \tau_b}{1 + K} \right) s + 1} \quad [11-8]$$

The pole locations are plotted in **Figure 11-8b**, with the locus of closed-loop poles shown for K increasing. Note the fundamental trade-off between high DC open loop gain (which means a small closed-loop DC error) and loop stability. For K approaching infinity, the closed-loop poles are very underdamped.

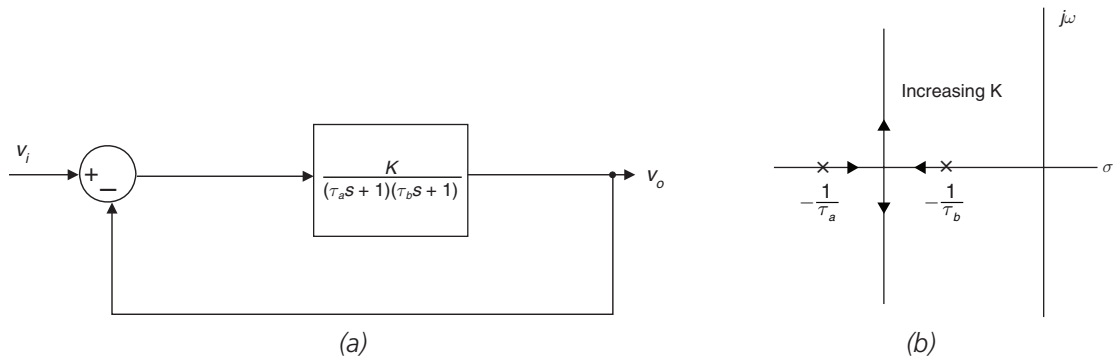


Figure 11-8: Second-order system with negative feedback loop.
(a) Block diagram. (b) Root locus as K increases.

Routh Stability Criterion

The Routh test is a mathematical test that can be used to determine how many roots of the characteristic equation lie in the right-half plane. When we use the Routh test, we don't calculate the location of the roots—rather we determine whether there are any roots at all in the right-half plane, without explicitly determining where they are.

The procedure for using the Routh test is as follows:

- Write the characteristic polynomial:

$$1 - L.T. = a_0 s^n + a_1 s^{n-1} + \cdots + a_n \quad [11-9]$$

Note that since we've written the characteristic polynomial $(1 - L.T.)$, we now are interested in finding whether there are *zeros* of $(1 - L.T.)$ in the right-half plane. Zeros of $(1 - L.T.)$ in the right-half plane correspond to closed-loop poles in the right-half plane. Furthermore, we assume that $a_n \neq 0$ for the analysis to proceed.

- Next, we see if any of the coefficients are zero or have a different sign from the others. A necessary (but not sufficient) condition for stability is that there are no nonzero coefficients in the characteristic equation and that all coefficients have the same sign.
- If all coefficients have the same sign, we next form a matrix with rows and columns in the following pattern, which is shown for n even.⁶ The table is filled horizontally and vertically until zeros are obtained in the rows. The third row and following rows are calculated from the previous two rows.

⁶ For n odd, a_n terminates the second row.

$$\begin{array}{cccccc}
 a_0 & a_2 & a_4 & \cdots & \cdots & \cdots \\
 a_1 & a_3 & a_5 & \cdots & \cdots & \cdots \\
 b_1 & b_2 & b_3 & \cdots & \cdots & \cdots \\
 c_1 & c_2 & c_3 & \cdots & \cdots & \cdots \\
 \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 0 & 0 & 0 & 0 & 0 & 0
 \end{array}$$

$$b_1 = \frac{- \begin{vmatrix} a_0 & a_2 \\ a_1 & a_3 \end{vmatrix}}{a_1} = \frac{a_1 a_2 - a_0 a_3}{a_1}$$

$$b_2 = \frac{- \begin{vmatrix} a_0 & a_4 \\ a_1 & a_5 \end{vmatrix}}{a_1} = \frac{a_1 a_4 - a_0 a_5}{a_1} \quad [11-10]$$

$$b_3 = \frac{- \begin{vmatrix} a_0 & a_6 \\ a_1 & a_7 \end{vmatrix}}{a_1} = \frac{a_1 a_6 - a_0 a_7}{a_1}$$

$$c_1 = \frac{- \begin{vmatrix} a_1 & a_3 \\ b_1 & b_2 \end{vmatrix}}{b_1} = \frac{b_1 a_3 - a_1 b_2}{b_1}$$

$$c_2 = \frac{- \begin{vmatrix} a_1 & a_5 \\ b_1 & b_3 \end{vmatrix}}{b_1} = \frac{a_5 b_1 - a_1 b_3}{b_1}$$

- The number of poles in the right-half plane is equal to the number of sign changes in the first column of the Routh matrix.

Let's apply the Routh test to the transfer function:

$$H(s) = \frac{1}{(s+1)(s+2)(s+3)(s-2)} = \frac{1}{s^4 + 4s^3 - s^2 - 16s - 12} \quad [11-11]$$

In this case, we already know that there is one right-half-plane pole at $s = +2$ radians/second, but we'll use the Routh test to verify this. The Routh matrix is:

$$\begin{array}{cccc}
 1 & -1 & -12 & 0 \\
 4 & -16 & 0 & 0 \\
 \left(\frac{-\begin{vmatrix} 1 & -1 \\ 4 & -16 \end{vmatrix}}{-1} = -12 \right) & \left(\frac{-\begin{vmatrix} 1 & -12 \\ 4 & 0 \end{vmatrix}}{-1} = 48 \right) & \left(\frac{-\begin{vmatrix} 1 & 0 \\ 4 & 0 \end{vmatrix}}{-1} = 0 \right) & 0 \\
 \left(\frac{-\begin{vmatrix} 4 & -16 \\ -12 & 48 \end{vmatrix}}{-12} = 0 \right) & \left(\frac{-\begin{vmatrix} 4 & 0 \\ -12 & 0 \end{vmatrix}}{-12} = 0 \right) & 0 & 0
 \end{array} \quad [11-12]$$

We see that there is one sign change in the first column, with the elements of the matrix changing from +4 to -12. Hence, there is one right-half-plane pole, as expected.

Let's next apply the Routh test to a system with three poles inside a unity-feedback loop (**Figure 11-9**). We'll use the Routh test to determine the values of K that result in stable operation of this feedback loop. The closed-loop transfer function for this system is:

$$\frac{v_o(s)}{v_i(s)} = \frac{\frac{K}{(s+1)^3}}{1 + \frac{K}{(s+1)^3}} = \left(\frac{K}{1+K} \right) \left(\frac{1}{\frac{s^3}{K+1} + \frac{3s^2}{K+1} + \frac{3s}{1+K} + 1} \right) \quad [11-13]$$

The denominator polynomial is:

$$D(s) = a_o s^3 + a_1 s^2 + a_2 s + a_3 = \left(\frac{1}{K+1} \right) s^3 + \left(\frac{3}{K+1} \right) s^2 + \left(\frac{3}{K+1} \right) s + 1 \quad [11-14]$$

The Routh matrix is:

$$\begin{array}{ccc}
 \left(\frac{1}{1+K} \right) & & \left(\frac{3}{1+K} \right) \\
 \left(\frac{3}{1+K} \right) & & 1 \\
 \frac{\left(\frac{3}{1+K} \right)^2 - \left(\frac{1}{1+K} \right)}{1+K} = \frac{8-K}{(1+K)^2} & & 0 \\
 1 & & 0 \\
 0 & & 0
 \end{array} \quad [11-15]$$

Note that if $K > 8$, there are two sign changes in the first column. Therefore, for $K = 8$, we expect two poles on the $j\omega$ axis and for $K > 8$ the system is unstable with two poles in the right-half plane. For $K < 8$, the system is stable with all three poles in the left-half plane.

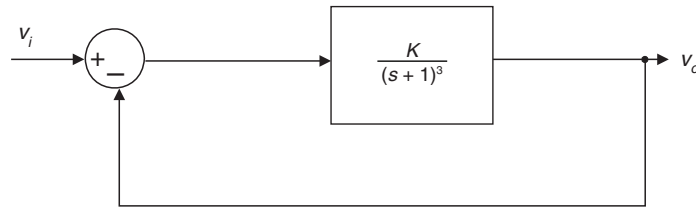


Figure 11-9: Three poles inside unity-feedback loop.

The Phase Margin and Gain Margin Tests

The previous analyses tell us what the bandwidth and DC gain of a closed-loop system is, but doesn't consider the question of whether a system will oscillate or have large amounts of overshoot. By using a simple Bode plot technique and a method known as the *phase margin method*, the relative stability of a feedback system can be determined. Phase margin is a very useful measure of the stability of a feedback system. The method for finding phase margin for a negative feedback system is as follows (**Figure 11-10**):

- Plot the magnitude and angle of the negative of the loop transmission, or $-a(s)f(s)$.
- Find the frequency where the magnitude of $a(s)f(s)$ drops to +1. This is the *crossover frequency* ω_c .
- The difference between the angle at the crossover frequency and -180° is the *phase margin* ϕ_m .
- The *gain margin* (G.M.) is defined as the change in open-loop gain required to make the system unstable. Systems with greater gain margins can withstand greater changes in system parameters before becoming unstable in closed loop.
- The *phase margin* is defined as the negative change in open-loop phase shift required to make a closed-loop system unstable.
- In general, a well-designed feedback loop has a phase margin of at least 45° and a gain margin (G.M.) > 3 or so.

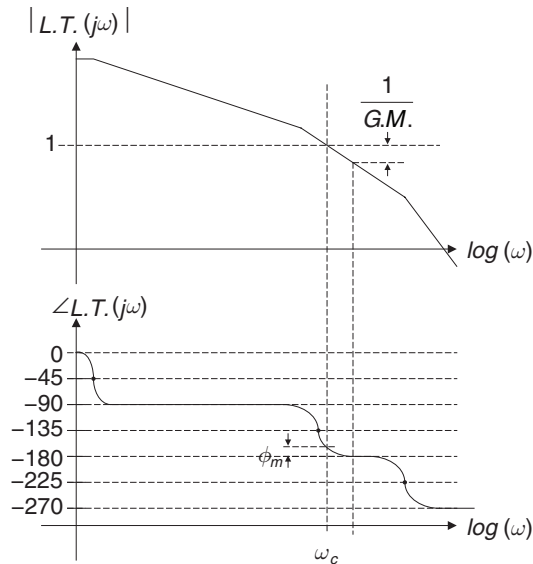


Figure 11-10: Plot of gain margin and phase margin.

Relationship Between Damping Ratio and Phase Margin

The damping ratio and phase margin are directly related. For a second-order system, a low phase margin in general implies a low damping ratio. For a standard second-order system with damping ratio < 0.6 , the relationship is approximately:

$$\zeta \approx \frac{\phi_m}{100} \quad [11-16]$$

Therefore, a damping ratio of 0.6 corresponds to a phase margin of 60° . The actual relationship over the range of damping ratios $0 < \zeta < 2$ is shown in **Figure 11-11**.

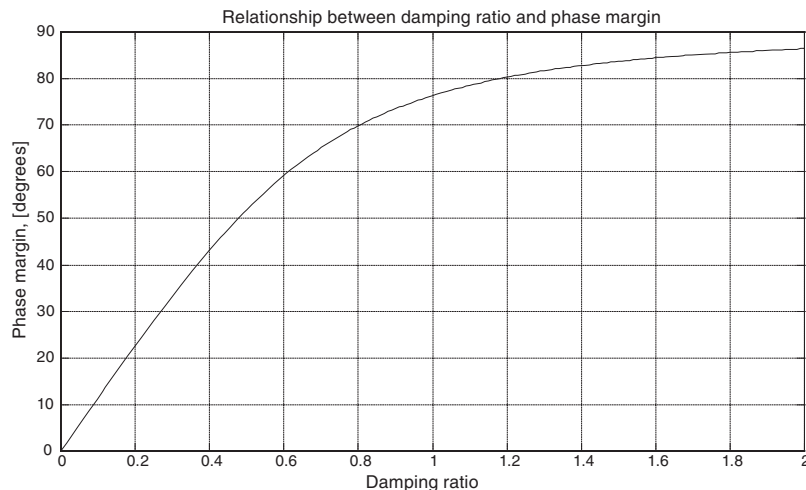


Figure 11-11: Relationship between phase margin and damping ratio.

Loop Compensation Techniques—Lead and Lag Networks

Several networks are available to compensate feedback networks. These networks can be added in series to the plant to modify the closed-loop transfer function, or placed in other locations in the feedback system. Shown next is a quick look at “lead” and “lag” networks.

The lag network (**Figure 11-12a**) is often used to reduce the gain of the loop transmission so that crossover occurs at a benign frequency. The transfer function of the lag network is:

$$H(s) = \frac{v_o(s)}{v_i(s)} = \frac{R_2Cs + 1}{(R_1 + R_2)Cs + 1} = \frac{\tau s + 1}{\alpha\tau s + 1}$$

$$\alpha = \frac{R_1 + R_2}{R_2} \quad [11-17]$$

$$\tau = R_2C$$

The Bode plot of the lag network (**Figure 11-12b**) shows that the network produces magnitude reduction at frequencies between the pole and the zero. When using a lag network, you’ll typically place the lag zero well below the crossover frequency of the loop. This ensures that the lag network doesn’t provide too much negative phase shift at crossover.

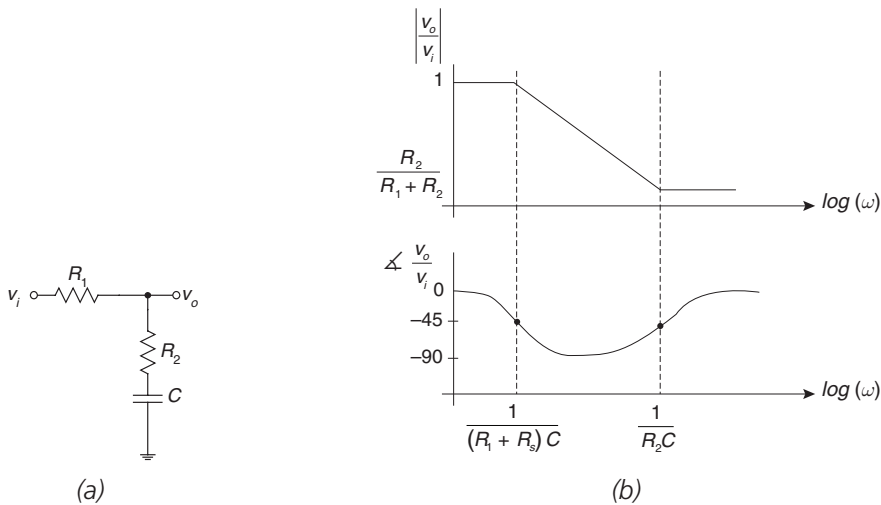


Figure 11-12: Lag network. (a) Circuit. (b) Bode plot of magnitude and phase angle of the frequency response of the lag network.

$$H(s) = \frac{v_o(s)}{v_i(s)} = \frac{R_2Cs + 1}{(R_1 + R_2)Cs + 1} = \frac{\tau s + 1}{\alpha\tau s + 1}$$

$$\alpha = \frac{R_1 + R_2}{R_2} \quad [11-18]$$

$$\tau = R_2C$$

The lead network (**Figure 11-13a**) is used to provide positive phase shift in the vicinity of the crossover frequency. The transfer function of the lead network is:

$$H(s) = \frac{v_o(s)}{v_i(s)} = \left(\frac{R_2}{R_1 + R_2} \right) \left(\frac{R_1 C s + 1}{\frac{R_1 R_2}{R_1 + R_2} C s + 1} \right) = \frac{1}{\alpha} \frac{\alpha \tau s + 1}{\tau s + 1} \quad [11-19]$$

$$\alpha = \frac{R_1 + R_2}{R_2}$$

$$\tau = (R_1 \parallel R_2) C$$

The Bode plot of the lead (**Figure 11-13b**) shows that the lead provided +45 degrees of positive phase shift at the zero frequency, while at the zero there is only +3dB of gain increase. When using a lead network, one generally places the lead zero near the crossover frequency of the loop to take advantage of the positive phase shift provided by the lead. The lead pole is then above crossover.

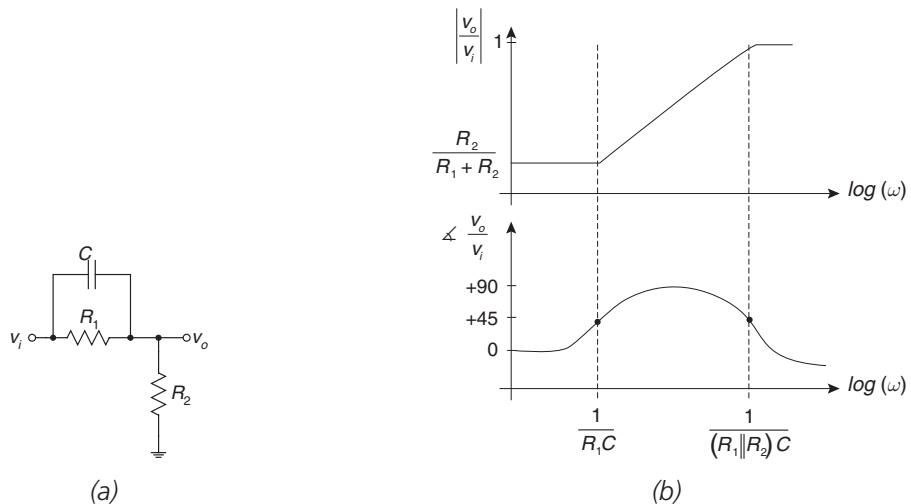


Figure 11-13: Lead network. (a) Circuit. (b) Bode plot of magnitude and phase angle of the frequency response of the lead network.

Parentetical Comment on Some Interesting Feedback Loops

The inquisitive student may wonder whether a system that has a loop transmission magnitude greater than unity where the loop transmission angle is -180° can be stable or not. In using the gain margin/phase margin test, we look at the frequency at which the magnitude drops to unity, and don't concern ourselves with other frequencies. By example, we'll show next that a system that has a loop transmission magnitude greater than unity where the loop transmission angle is -180° can be stable. It's understood that this is not necessarily an intuitive result, but

Chapter 11

we'll run with it anyway. Consider the system of **Figure 11-14a**, which is a unity-feedback system with two zeros and three poles in the forward path.

The negative of the loop transmission for this system is:

$$-L.T. = \frac{100(s+1)^2}{s^3} \quad [11-20]$$

A plot of the negative of the loop transmission is shown in **Figure 11-14b**. Note that the loop transmission magnitude is greater than unity at the frequency when the angle of the negative of the loop transmission is -180° . In this case, the angle is less than -180° up to 1 radian/second.

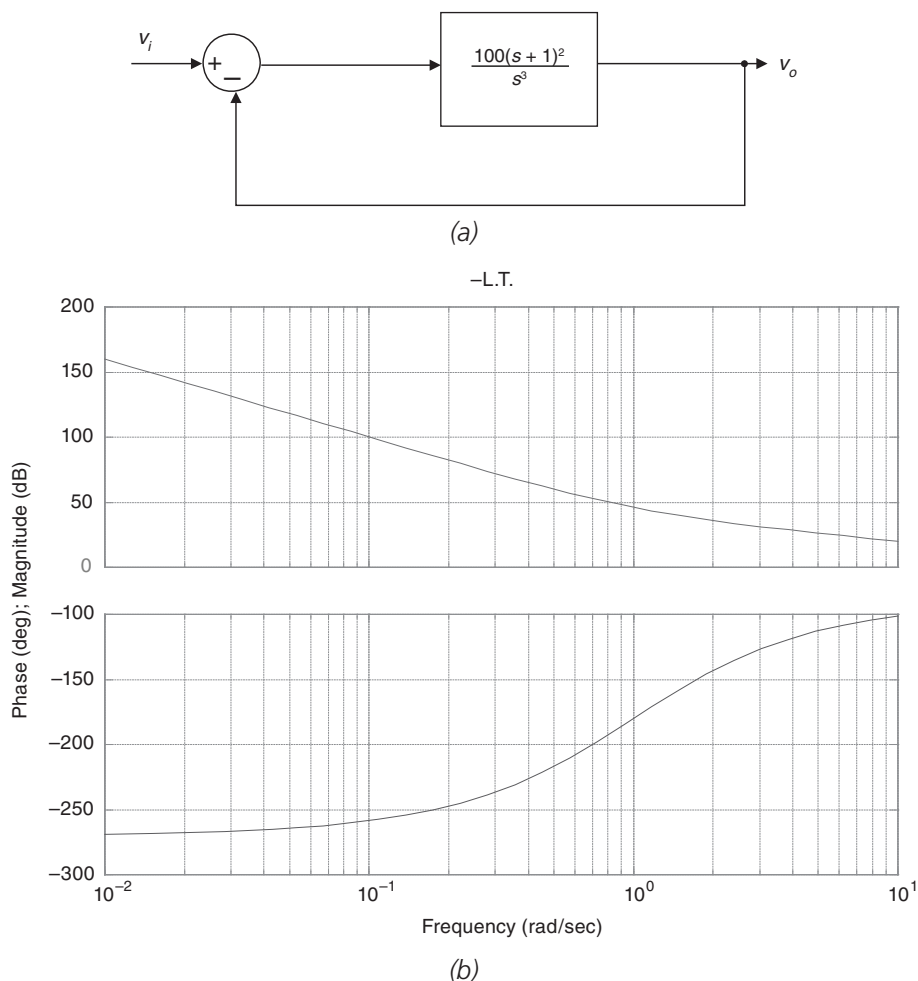


Figure 11-14: Unity-feedback system that has a loop transmission magnitude greater than unity where the loop transmission angle is -180° . (a) System. (b) Bode plot of $-L.T.$ showing the magnitude is greater than 1 when the angle is more negative than -180° .

We'll next use the Routh test to determine the stability of this system. The closed-loop transfer function for this system is:

$$\frac{v_o(s)}{v_i(s)} = \frac{\frac{100(s+1)^2}{s^3}}{1 + \frac{100(s+1)^2}{s^3}} = \frac{100(s^2 + 2s + 1)}{s^3 + 100s^2 + 200s + 100} \quad [11-21]$$

We can use the Routh criteria or we can factor the denominator of this transfer function to determine stability. The denominator polynomial is:

$$D(s) = s^3 + 100s^2 + 200s + 100 \quad [11-22]$$

The Routh matrix is as follows:

$$\begin{array}{cc} 1 & 200 \\ 100 & 100 \\ \frac{(100)(200) - (1)(100)}{100} = 199 & 0 \\ \frac{(199)(100) - (100)(0)}{199} = 100 & 0 \\ 0 & 0 \end{array} \quad [11-23]$$

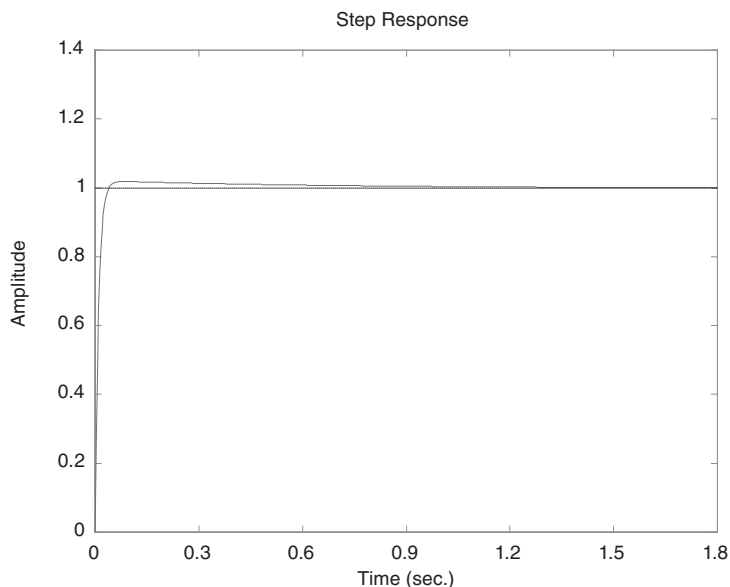
The Routh test shows that this system is BIBO stable since there are no sign changes in the first column of the Routh matrix. Numerical analysis shows that the closed-loop poles and zeros are at frequencies:

Zeros:	2 zeros at -1 r/s
Poles:	Poles at -97.97 r/s, -1.12 r/s and -0.92 r/s

Therefore, all poles are in the left-half plane and the system is BIBO stable. The closed-loop step response (**Figure 11-15**) confirms that the system is stable. Note the long decaying “tail” of the step response while it settles to unity gain. This long tail is characteristic of systems with closely spaced poles and zeros.⁷

⁷ Also known as a pole/zero “doublet.”

Figure 11-15: Step response of unity-feedback system that has a loop transmission magnitude greater than unity where the loop transmission angle is -180° .



Example 11.1: Gain of +1 amplifier

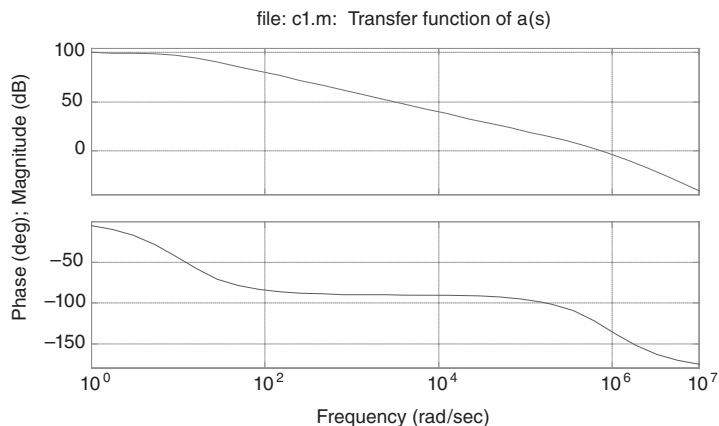
Consider an operational amplifier with a DC gain of 10^5 , a low-frequency pole at 10 radians/second and a high-frequency pole at 10^6 radians/second. This transfer function $a(s)$ is representative of many commercially available operational amplifiers, and is expressed as:

$$a(s) = \frac{10^5}{(0.1s + 1)(10^{-6}s + 1)} \quad [11-24]$$

What is the bandwidth and risetime when this op-amp is configured as a gain of +1 amplifier?

Shown in **Figure 11-16** is the Bode plot of the forward path gain $a(s)$. Note that the DC gain is 10^5 (100dB) and the phase starts at 0° and falls asymptotically to -180° at frequencies much higher than 10^6 radians per second (above the second pole).

Figure 11-16: Open-loop transfer function of $a(s)$ for Examples 11.1 and 11.2.



The gain of +1 op-amp circuit is shown in **Figure 11-17a** where the op-amp has unity feedback. The block diagram of this circuit is shown in **Figure 11-17b**.

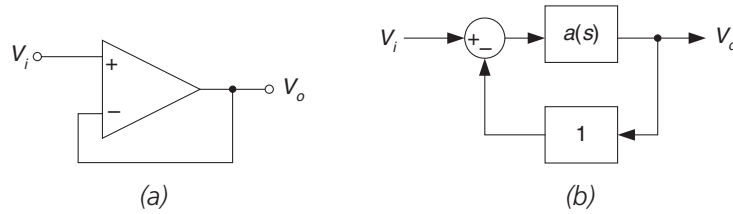


Figure 11-17: Gain of +1 op-amp circuit. (a) Circuit. (b) Block diagram.

For this circuit, the negative of the loop transmission is:

$$-L.T. = a(s)f = \frac{10^5}{(0.1s + 1)(10^{-6}s + 1)} \quad [11-25]$$

which is the same as the open-loop transfer function of the op-amp $a(s)$. Using MATLAB, the bandwidth and phase margin are calculated, as shown in **Figure 11-18**. The results show a phase margin of 52° and a crossover frequency of 786,150 radians/second (125 kHz). From this analysis, we expect some overshoot in the step response (since the phase margin results in a damping ratio of ~ 0.5), some overshoot in the frequency response, and a 10–90% risetime⁸ of approximately $0.35/125000 = 2.8 \mu\text{s}$ (**Figure 11-19**).

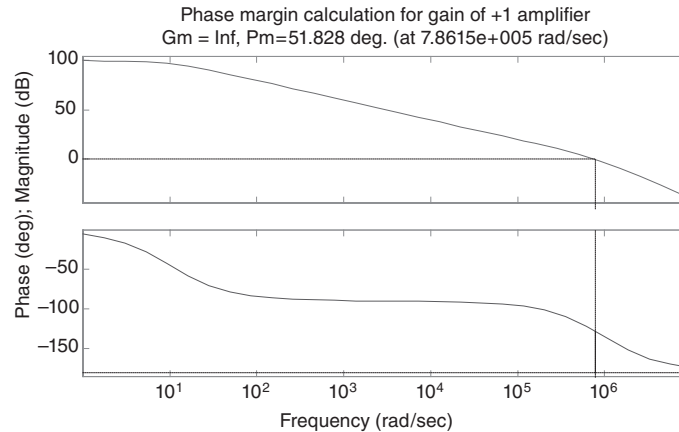


Figure 11-18: Phase margin and bandwidth calculation for gain of +1 op-amp circuit.

⁸ In general, an estimate of 10–90% risetime is $0.35/f_c$, where f_c is the crossover frequency in Hz.

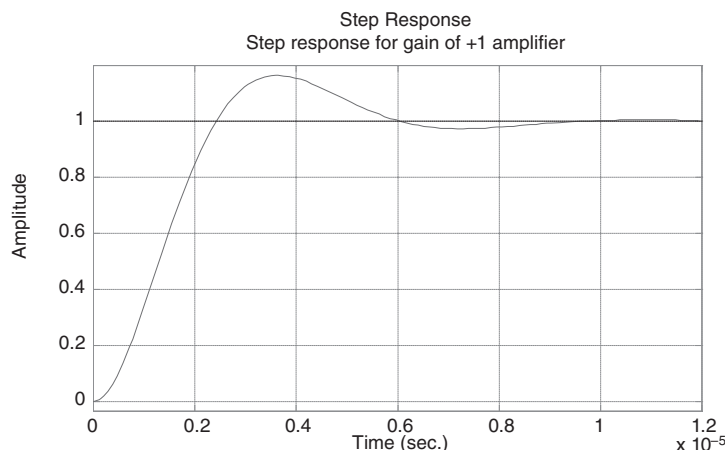


Figure 11-19: Step response for gain of +1 amplifier.

Example 11.2: Gain of +10 amplifier

What is the bandwidth and risetime when this same op-amp is configured as a gain of +10 amplifier? The gain of +10 op-amp circuit is shown in **Figure 11-20a**, and the block diagram is shown in **Figure 11-20b**. Note that the $9R/R$ divider gives a feedback factor $f = 0.1$.

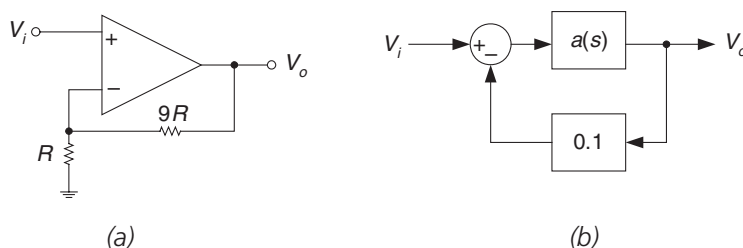


Figure 11-20: Gain of +10 op-amp circuit. (a) Circuit. (b) Block diagram.

For this circuit, the negative of the loop transmission is:

$$-L.T. = a(s)f = \frac{10^4}{(0.1s + 1)(10^{-6}s + 1)} \quad [11-26]$$

When we plot the loop transmission magnitude and phase (**Figure 11-21**) the results show a phase margin of 84° and a crossover frequency of 99,509 radians/second (15.8 kHz). From this analysis, we expect no overshoot in the step response and frequency response (since the phase margin is close to 90°), and a 10–90% risetime of approximately $0.35/15,800 = 22 \mu\text{s}$, as shown in **Figure 11-22**. Therefore the system is well damped and there is no overshoot in the step response (**Figure 11-23**).

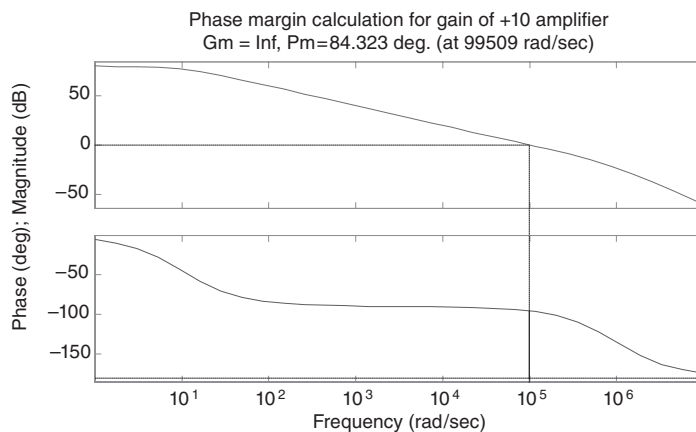


Figure 11-21: Phase margin and bandwidth calculation for gain of +10 op-amp circuit.

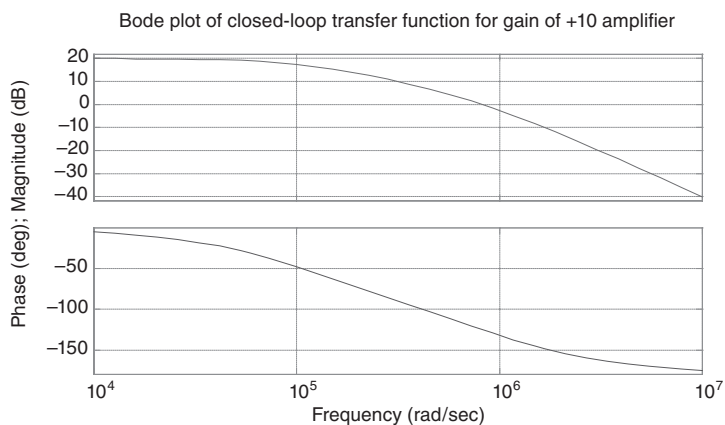


Figure 11-22: Bode plot of closed-loop transfer function for gain of +10 amplifier.

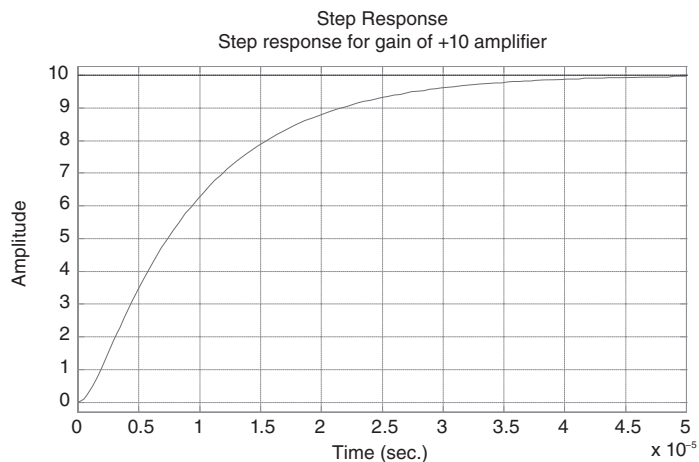


Figure 11-23: Step response for gain of +10 amplifier.

Example 11.3: Integral control of reactive load

Consider the op-amp driving a reactive load as shown in **Figure 11-24**. Assume that the op-amp is ideal; it has infinite bandwidth and can source and sink infinite current. Given this, the transfer function is:

$$H(s) = \frac{v_o(s)}{v_i(s)} = \frac{1}{LCs^2 + \frac{L}{R}s + 1} = \frac{1}{10^{-10}s^2 + 10^{-6}s + 1} \quad [11-27]$$

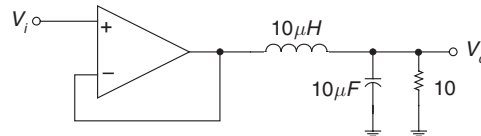


Figure 11-24: Ideal op-amp driving a reactive load.

This second-order reactive load $H(s)$ has the following:

- Poles: $-5 \times 10^3 \pm 9.99 \times 10^4 j$
- Damping ratio: 0.05
- $Z_o = 1 \Omega$
- Q : 10

When plotting this transfer function $H(s)$, we see the underdamped response of **Figure 11-25**.

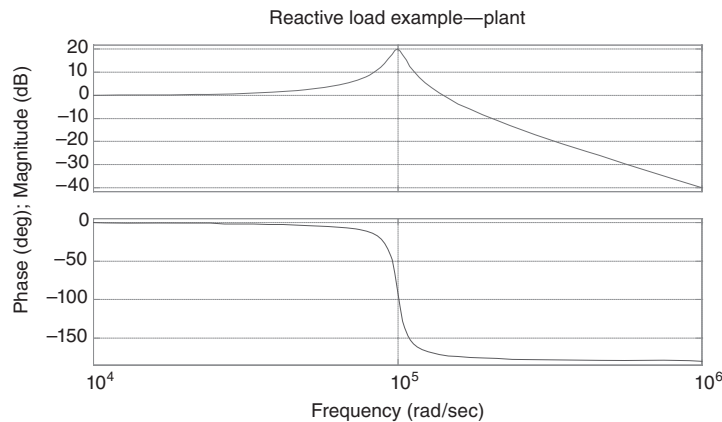


Figure 11-25: Reactive load example plant Bode plot.

In this example, we'll design a closed-loop controller to regulate the output voltage (**Figure 11-26**). We'll assume that the forward path compensator $G_c(s)$ includes an integrator, so that there will be zero DC error⁹ in V_o .

⁹ Having an integrator in $G_c(s)$ ensures that the DC error of the loop is zero. If there were an error, the integrator would forever integrate to infinity. Hence, there must be zero DC error.

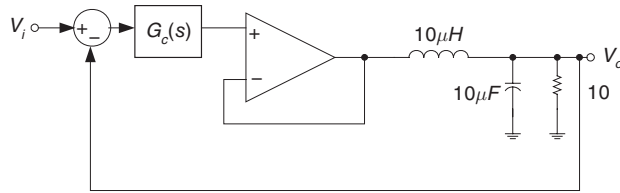


Figure 11-26: Closed loop controller for Example 11.3.

As a first attempt, let's try a compensator transfer function of the form:

$$G_c(s) = \frac{4 \times 10^3}{s} \quad [11-28]$$

The resultant loop transmission is:

$$-L.T. = \left(\frac{1}{10^{-10}s^2 + 10^{-6}s + 1} \right) \left(\frac{4 \times 10^3}{s} \right) \quad [11-29]$$

Plotting this loop transmission magnitude and phase yields some interesting results, as shown in **Figure 11-27**.

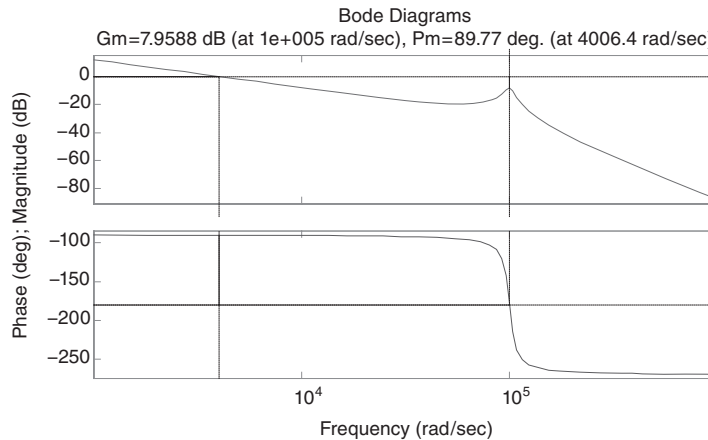


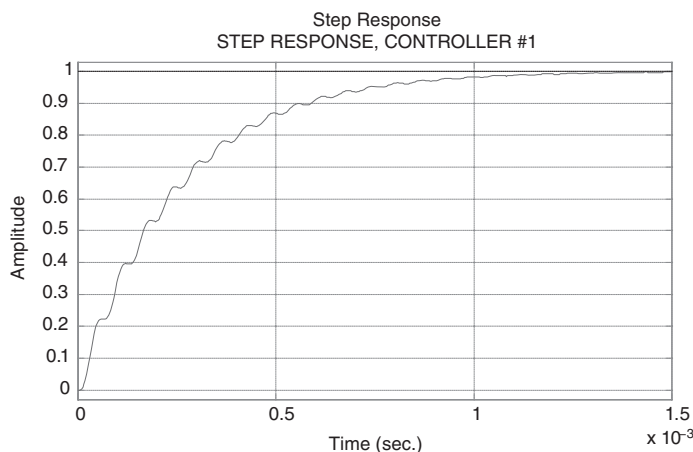
Figure 11-27: Plot of magnitude and phase of the loop transmission, Attempt #1, showing a crossover frequency of 4006 radians/second and a phase margin of 89.8°.

Results are:

- Crossover frequency: 4006 radians/sec (637 Hz)
- Phase margin: 90°
- Given this, we expect a well-controlled step response, with risetime ~0.5 ms.

The resultant step response of the closed-loop system is shown in **Figure 11-28**.

Figure 11-28: Step response, Attempt #1.



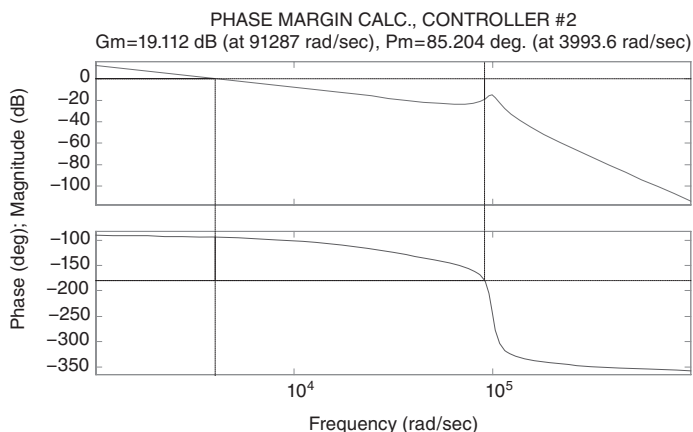
Things look OK except for the oscillatory behavior on the rising edge. What's going on here? The Bode plot of the loop transmission phase and magnitude tells the story. Although the phase margin is fine, the gain margin isn't so great, due to the underdamped pole pair. In order to help this oscillation problem, we might attempt to add a pole above crossover to damp out the underdamped pair. Let's try adding a pole at 5×10^4 radians/second to damp the complex poles at 10^5 radians/sec. This results in:

$$G_c(s) = \left(\frac{4 \times 10^3}{s} \right) \left(\frac{1}{2 \times 10^{-5} s + 1} \right) \quad [11-30]$$

The results of a MATLAB crossover frequency and phase margin test are shown in **Figure 11-29**. Note that, by adding the low-pass filter, we've significantly improved the gain margin of this circuit. Results show:

- Crossover frequency: 3993 radians/sec (635 Hz)
- Phase margin: 85°
- Better-behaved step response, as shown in **Figure 11-30**.

Figure 11-29: Plot of the loop transmission magnitude and phase, Attempt #2, showing crossover frequency of 3993 radians/second and a phase margin of 85° .



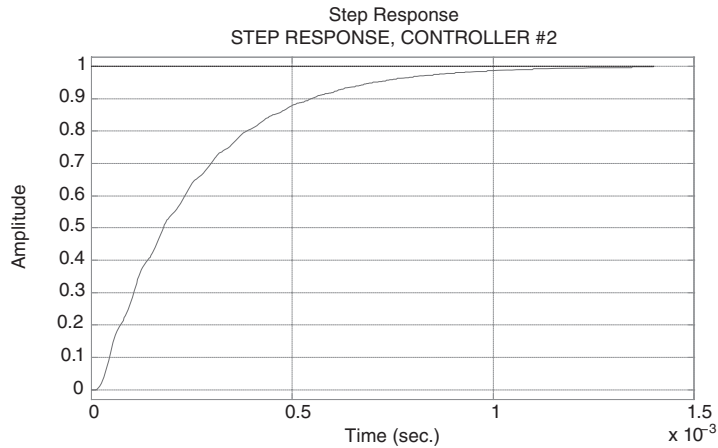


Figure 11-30: Step response, Attempt #2.

Example 11.4: Photodiode amplifier

A photodiode amplifier is shown in **Figure 11-31a**. The photodiode puts out a current proportional to the light hitting it, and the *transimpedance* connection of the op-amp converts this photodiode current into an output voltage. The input-to-output ideal transfer function is:

$$\frac{v_o}{i_p} = -R_f \quad [11-31]$$

When modeling this circuit, the photodiode is modeled as a current source in parallel with a parasitic capacitance C_p , as shown in **Figure 11-31b**.

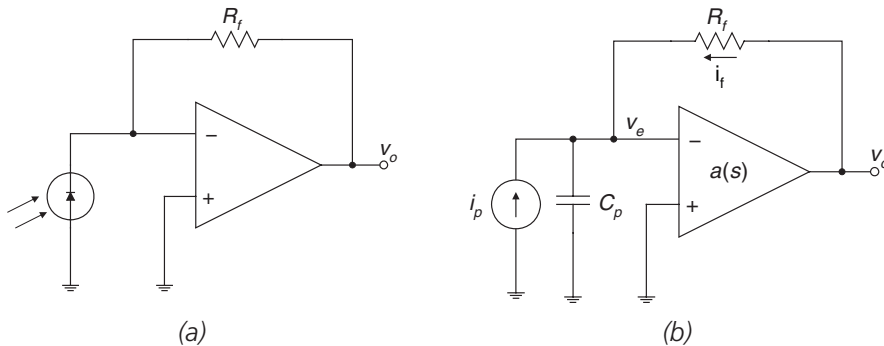
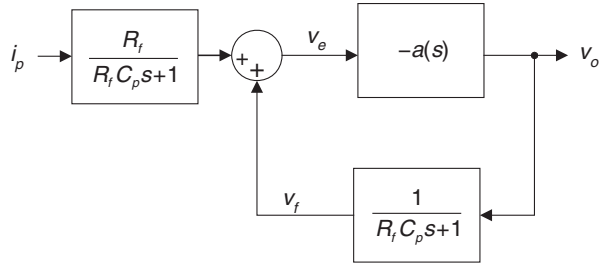


Figure 11-31: Photodiode amplifier. (a) Circuit. (b) Model, showing equivalent model of photodiode that includes a current source i_p and parasitic capacitance C_p .

The block diagram for this system is shown in **Figure 11-32**.

Figure 11-32: Block diagram of photodiode amplifier. The block $a(s)$ is the gain of the operational amplifier.



The closed-loop transfer function of this system is:

$$\frac{v_o}{i_p} = -\frac{R_f}{R_f C_p s + 1} \left(\frac{a(s)}{1 + \frac{a(s)}{R_f C_p s + 1}} \right) \quad [11-32]$$

Note that if the loop transmission is much larger than 1 the approximate transfer function is $-R_f$. The loop transmission of this system, easily found by inspection, is:

$$-L.T. = \frac{a(s)}{R_f C_p s + 1} \quad [11-33]$$

In general, an operational amplifier will have a dominant pole and a second pole near the crossover frequency, resulting in:

$$-L.T. = \frac{a_o}{s(\tau s + 1)(R_f C_p s + 1)} \quad [11-34]$$

This means that there are three poles (at least) in the loop transmission. If we attempt to close a feedback loop with a bandwidth greater than $1/R_f C_p$ there are potential problems with stability.

A model of an actual system, using the CLC426 opamp, was created, including dominant pole, second pole, and output resistance.

Parasitics in this circuit are as follows:

- L_{cable} : Inductance of cable connecting photodiode to PC board. Approximate inductance is 10 nanohenries per centimeter of length. A value of 50 nanohenries was used for all simulations.
- L_{pr} : Inductance in series with feedback resistance; approximately 10 nH.
- L_{pc} : Inductance in series with feedback capacitor; approximately 5 nH.
- C_p : Capacitance of photodiode. Approximately 60 pF.
- C_{in} : Input capacitance of op-amp, approximately 5 pF.

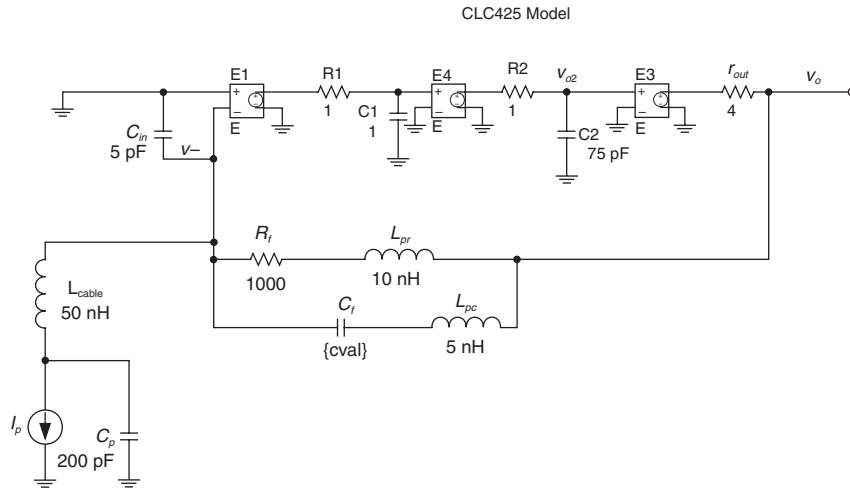


Figure 11-33: Photodiode amplifier model

The $R_f C_p$ combination results in a pole inside the feedback loop that results in potential instability. Plotted in **Figure 11-34** is frequency response of v_o/i_p , of the original circuit with $R_f = 15k$, showing that the response is very underdamped. Other unmodeled poles can result in oscillation.

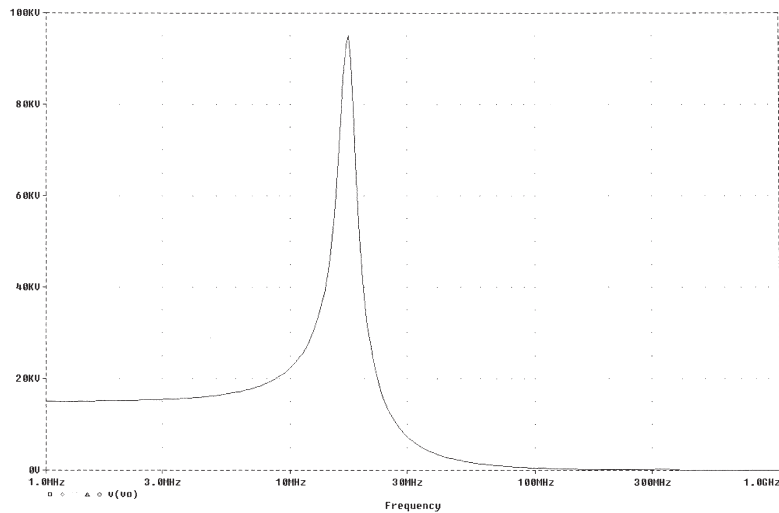


Figure 11-34: PSpice results of frequency response of the photodiode amplifier, showing potential instability near 15 MHz.

By adding a capacitance across the feedback resistor, a *lead* transfer function is created. The added lead zero can create positive phase shift near the crossover frequency, hence improving stability. The feedback factor becomes:

$$f(s) = \frac{\frac{1}{C_p s}}{\frac{1}{C_p s} + \frac{R_f}{R_f C_f s + 1}} = \frac{1}{1 + \frac{R_f C_p s}{R_f C_f s + 1}} = \frac{R_f C_f s + 1}{R_f (C_f + C_p) s + 1} \quad [11-35]$$

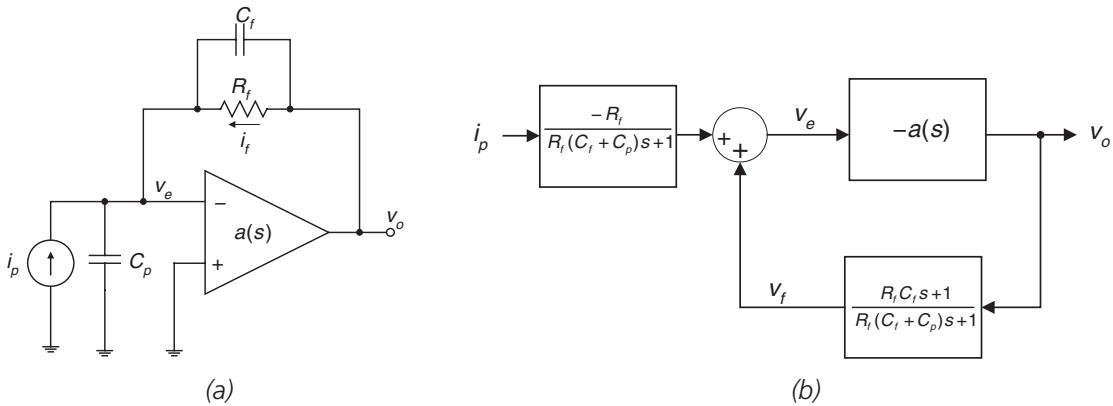


Figure 11-35: Photodiode amplifier with lead compensation. (a) Circuit. (b) Block diagram.

Simulations were modified for $R_f = 1\text{ k}$ and C_f adjustable from 2 pF to 10 pF. Results show (**Figure 11-36**) that it may be possible to achieve ~50 MHz bandwidth by proper adjustment of C_f . This, of course, depends on the accuracy of these simulations. Measurements should be made on the prototype to verify/refute these models.

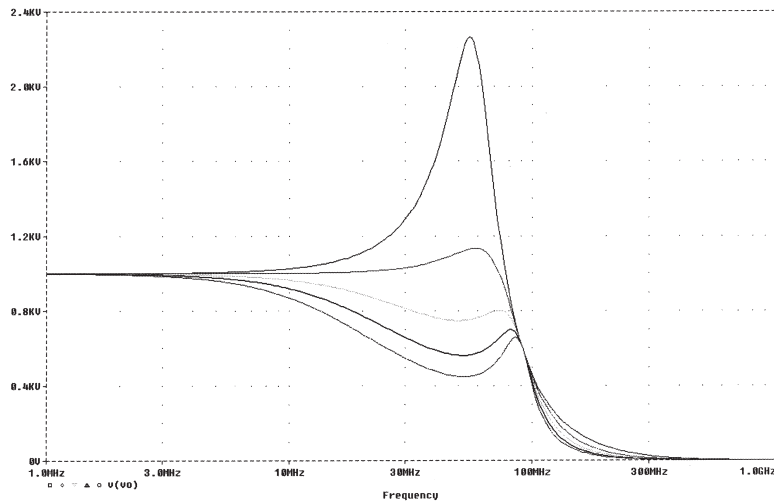


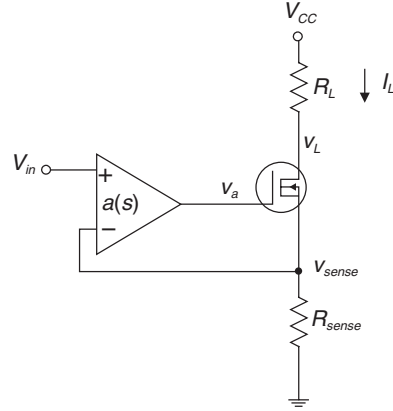
Figure 11-36: Frequency response results for photodiode amplifier with $C_f = 2, 4, 6, 8, 10\text{ pF}$, with $R_f = 1\text{ k}$ and $C_p = 60\text{ pF}$.

Example 11.5: MOSFET current source

In **Figure 11-37** is a MOSFET current source, with an op-amp used in a negative feedback configuration to maintain control of the MOSFET drain current. If the op-amp is ideal and if the MOSFET is operating in the linear region, the input-output transfer function is:

$$\frac{I_L}{V_{in}} \approx \frac{1}{R_{sense}} \quad [11-36]$$

Figure 11-37: MOSFET current source. The voltage V_{sense} senses MOSFET drain current.



This result is contingent on the feedback control system being stable (i.e., not oscillating). A small-signal model is shown in **Figure 11-38**. Following are the parameters:

- τ_h Time constant of op-amp high-frequency pole
- r_{out} Output resistance of op-amp
- C_{gs} MOSFET gate-source capacitance
- C_{gd} MOSFET gate-drain capacitance
- g_m MOSFET transconductance
- R_L Load resistance

The transfer function of the MOSFET source follower is estimated using the method of open-circuit time constants. This method assumes that a single pole dominates the transfer function; the resultant transfer function from the output of the op-amp to v_{sense} is:

$$\begin{aligned} \frac{v_{sense}}{v_a} &\approx \frac{A_o}{\tau s + 1} \\ A_o &= \frac{g_m R_{sense}}{1 + g_m R_{sense}} \\ \tau &= \left[\frac{r_{out} + R_{sense}}{1 + g_m R_{sense}} \right] C_{gs} + [r_{out} + R_L + G_M r_{out} R_L] C_{gd} \\ G_M &= \frac{g_m}{1 + g_m R_{sense}} \end{aligned} \quad [11-37]$$

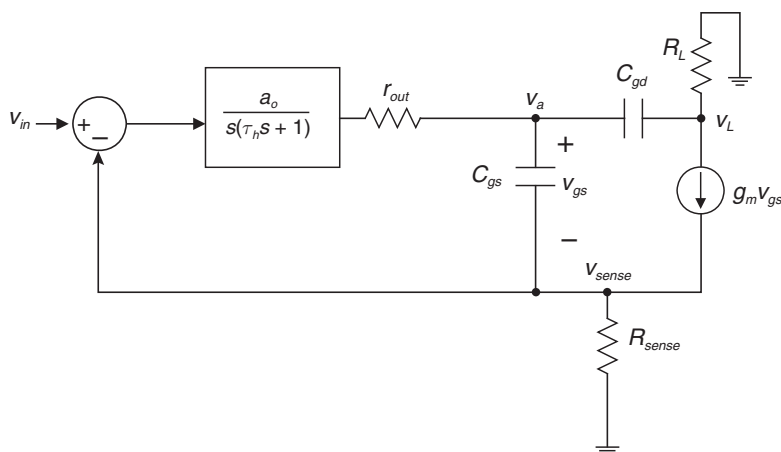


Figure 11-38: MOSFET current source—small-signal model.

The system was modeled assuming a TLO84 op-amp (with gain-bandwidth product of 4 MHz) and a IRF7403 MOSFET with $g_m = 10$ A/V, $C_{gs} = 1040$ pF and $C_{gd} = 160$ pF and a load resistance $R_L = 1\Omega$. In **Figure 11-39** is 6.7 Mrad/second and a phase margin of 28° . Due to the low phase margin, significant overshoot in the step response is expected. This is shown in **Figure 11-40**.

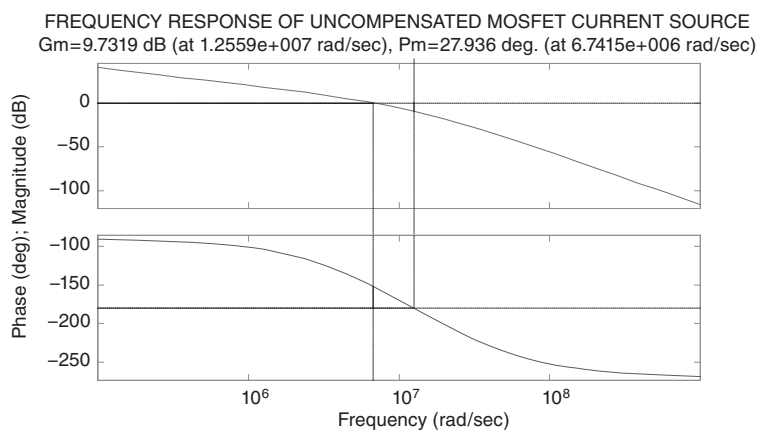
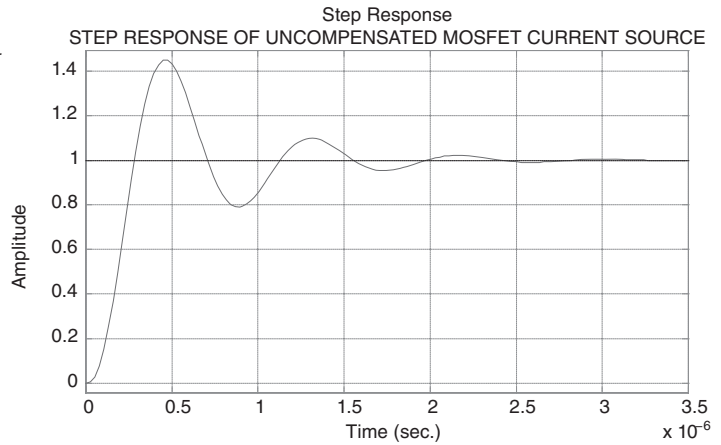


Figure 11-39: MOSFET current source—Bode plot of loop transmission showing crossover frequency of 6.7 Mrad/sec and a phase margin of 28° .

Figure 11-40: MOSFET current source—step response.



By adding lag compensation (**Figure 11-41a**) the system can be stabilized to provide more phase margin. Shown in **Figure 11-41b** is the block diagram of the system. The lag compensation adds a zero in the loop transmission at $-R_i C$ and a lag pole at $-(R_f + R_i)C$.

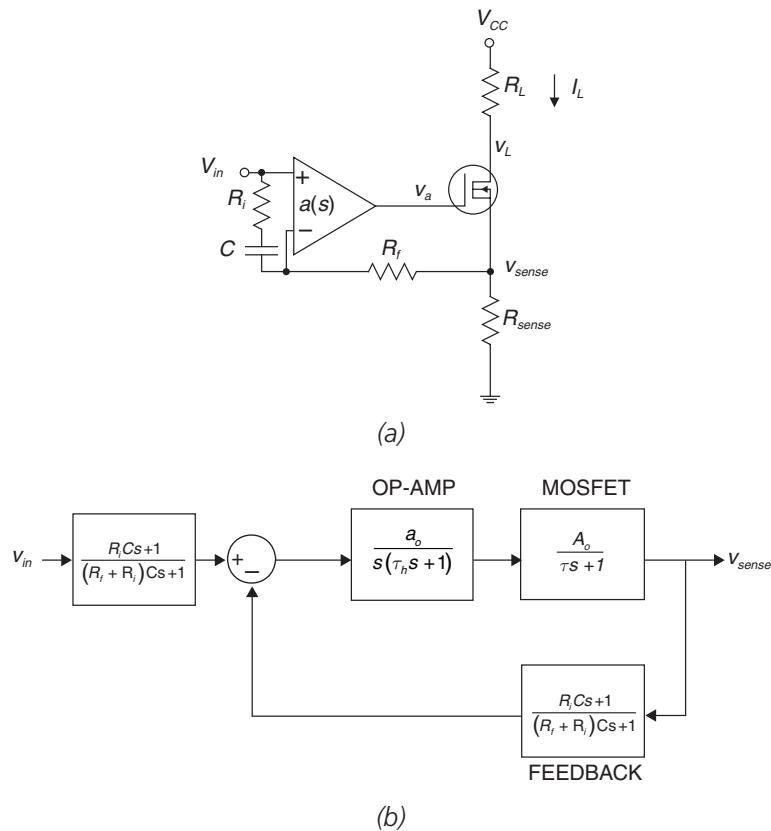


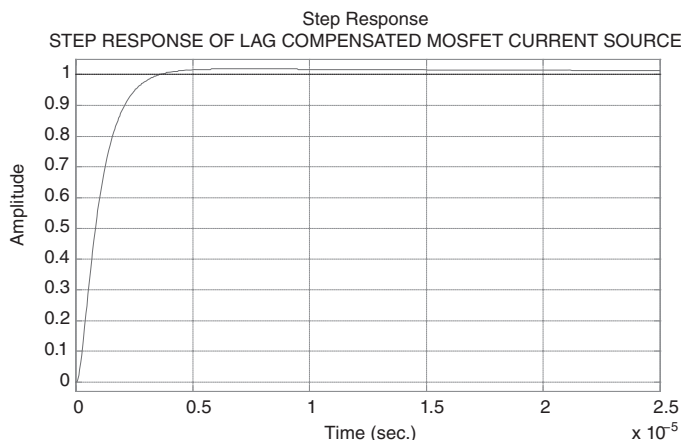
Figure 11-41: MOSFET current source with lag compensation
(a) Circuit with lag components R_i , C and R_f added. (b) Block diagram.

Step response results for the resultant system with the following parameters are shown in **Figure 11-42**.

- $R_i = 47k$
- $R_f = 470k$
- $C = 1000$ pF

Note that the system has a better-behaved step response, with no overshoot, but that the 10–90% risetime has significantly increased as compared to the uncompensated case.

Figure 11-42: MOSFET current source with lag compensation, step response.



Example 11.6: Maglev example

Maglev¹⁰ systems using superconducting magnets in the magnetic suspension have low damping. Furthermore, it has been demonstrated that these electrodynamic suspensions (EDS) may have slightly negative damping under certain operating conditions (with poles in the right-half plane). Therefore, a control system is needed to prevent underdamped or unstable vertical oscillations.

In order to suspend the magnet statically, the downward gravitational pull is canceled by an upward magnetic force. The magnet is levitated stably; that is, a deviation from the equilibrium position results in a restoring force, similar to a mass and a spring. In an electromagnetic suspension (such as a steel ball suspended in a magnetic field) there is no stable equilibrium for DC excitation of the magnet.

The magnetic levitating force acting on the magnet is given by:

$$f_z = -k_m z = -C i_M^2 z \quad [11-38]$$

where k_m is the equivalent spring constant, z is the vertical distance with reference to the magnet null position, i_M is magnet current, and C is a constant that accounts for magnet and coil geometry and relative velocity between the magnet and levitating coils.

¹⁰ “Maglev” is a term used generically for a number of systems utilizing magnetic suspensions for ground transportation. Currently (2004) there is one revenue-producing Maglev system operating in China, and others are proposed.

Assuming that there are incremental changes in forces, magnet vertical position, and magnet currents, a linearized model can be generated relating incremental changes in magnet vertical position to changes in incremental magnet current. Vertical force, vertical position and magnet current are given as the sum of a DC component and an incremental component:

$$\begin{aligned} f_z &= F_z + \tilde{f}_z \\ z &= Z_o + \tilde{z} \\ i_M &= I_M + \tilde{i}_m \end{aligned} \quad [11-39]$$

Putting this into the force equation results in:

$$f_z \approx -CI_M^2 Z_o - CI_M^2 \tilde{z} - 2CI_M Z_o \tilde{i}_m \quad [11-40]$$

where second-order and higher terms have been neglected. At equilibrium, there is a resultant magnetic force that balances the force of gravity:

$$F_z = Mg = -CI_M^2 Z_o \quad [11-41]$$

Newton's law applied to the magnet results in:

$$M \frac{d^2 \tilde{z}}{dt^2} = f_z - Mg = -CI_M^2 \tilde{z} - 2CI_M Z_o \tilde{i}_m \quad [11-42]$$

resulting in:

$$\frac{M}{CI_M^2} \frac{d^2 \tilde{z}}{dt^2} + \tilde{z} = \frac{2Mg}{k_m I_M} \tilde{i}_m \quad [11-43]$$

Using the spring constant k , and converting the equation to the frequency domain results in:

$$\left(\frac{M}{k_m} s^2 + 1 \right) z(s) = \frac{2Mg}{k_m I_M} i_m(s) \quad [11-44]$$

resulting in the transfer function between magnet position and magnet control current:

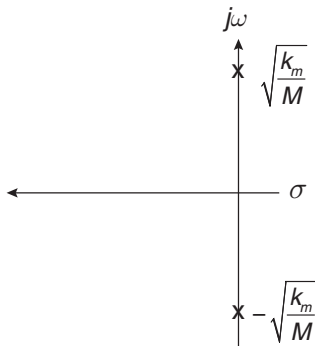
$$\frac{z(s)}{i_m(s)} = \frac{2Mg}{k_m I_M \left(\frac{M}{k_m} s^2 + 1 \right)} \quad [11-45]$$

This result shows that this suspension has two $j\omega$ -axis poles, as in a simple, lossless mass-spring system:

$$s_{p1,2} = \pm j \sqrt{\frac{k_m}{M}} \quad [11-46]$$

Such a system can have the closed-loop poles arbitrarily adjusted by applying position-velocity. By adjusting the parameters K_p , K_v , and K_p the poles can be placed in the left-half plane with sufficient damping to achieve good ride quality.

Figure 11-43: Pole plot of plant for Maglev example, showing poles on the $j\omega$ axis.



Representative numbers for a Maglev suspension magnet section are as follows:

$$M = 10,000 \text{ kg}$$

$$k_m = 10^5 \text{ Newtons/cm} = 10^7 \text{ Newtons/meter}$$

$$I_M = 10^4 \text{ amps}$$

This results in a resonant frequency of $\omega_o = 31.6 \text{ r/s}$ (with $f_o = 5 \text{ Hz}$), and a plant transfer function:

$$a(s) = \frac{z(s)}{i_m(s)} = \frac{1.96 \times 10^{-6}}{(10^{-3}s^2 + 1)} = \frac{A}{\frac{s^2}{\omega_o^2} + 1} \quad [11-47]$$

The poles are on the $j\omega$ axis, corresponding to the underdamped suspension (**Figure 11-44**). In order to improve ride quality for the passengers, the suspension poles must be moved into the left-half plane by selection of suitable compensation.

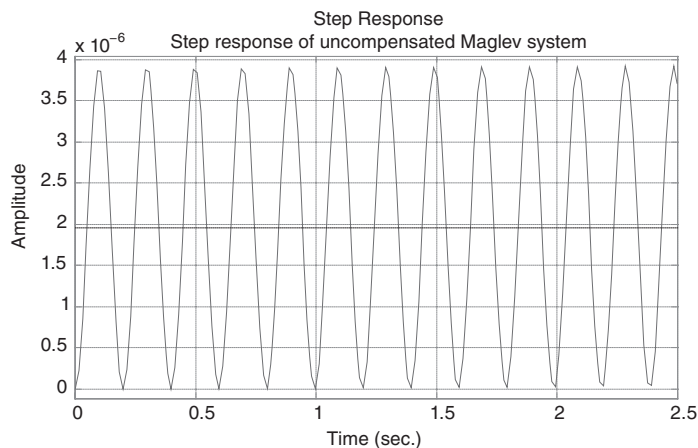


Figure 11-44: Step response of uncompensated Maglev suspension.

The system is compensated using velocity feedback, as shown in **Figure 11-45**. Velocity feedback is equivalent to adding damping to the system.

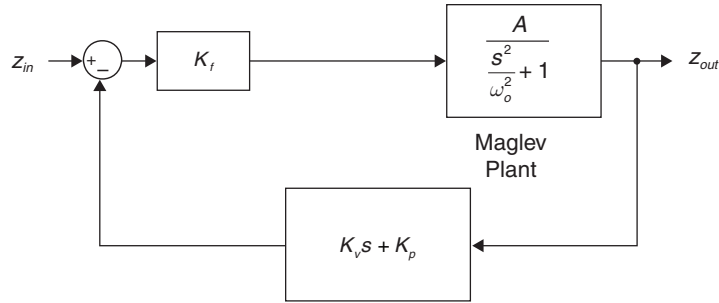


Figure 11-45: Control system block diagram for Maglev example.

A control system with $K_v = 10^5$ and $K_p = 10^4$ results in a system with closed-loop transfer function:

$$H(s) = \frac{1.64 \times 10^{-6}}{8.36 \times 10^{-4} s^2 + 1.64 \times 10^{-2} s + 1} \quad [11-48]$$

The closed-loop poles have a damping ratio of $\zeta = 0.28$ and at pole locations $-9.8 \pm j(33.2)$ radians/second. Therefore, we expect some oscillation near 33 radians/second (5.2 Hz). The resultant control system results in a much more well-behaved step response (**Figure 11-46**).

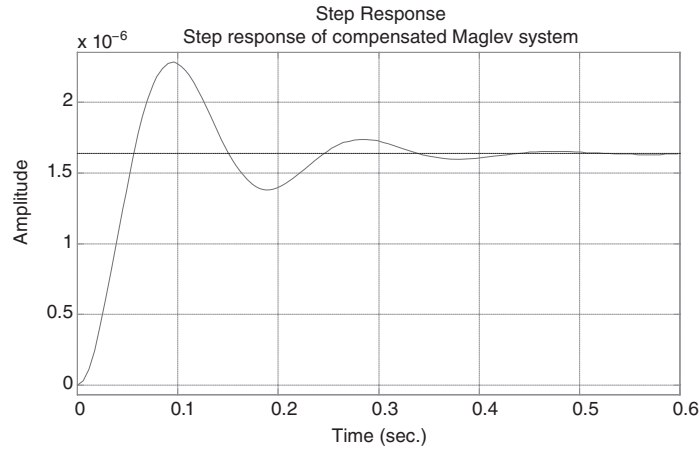


Figure 11-46: Step response of compensated Maglev system.

Appendix: MATLAB Scripts

MATLAB script for gain of +1 and +10 amplifiers

```
function c1
% Control system example #1
% Calculates parameters for gain of +1 and gain of +10 amplifiers
% Marc Thompson, 10/22/99

% Open loop transfer function a(s)
ao=1e5; % DC gain
d=conv([0.1 1],[1e-6 1]); % Poles
a=tf(ao,d); % Create transfer function a(s)
bode(a) % Plot Bode plot of a(s)
title('file: c1.m: Transfer function of a(s)')
figure;
pzmap(a); % Plot pole/zero map of a(s)
title('file: c1.m: Pole map of a(s)');
damp(a) % Find natural frequency and damping ratio of
a(s)

% gain of +1
f=1; % Feedback gain of +1
margin(a*f) % Find phase and gain margin
title('Phase margin calculation for gain of +1 amplifier');figure;
f=tf(f,1); % Create feedback f(s)
cloop=feedback(a,f) % Close the loop, find transfer function H(s)
bode(cloop)
title('Bode plot of closed-loop transfer function for gain of +1
amplifier');figure
step(cloop)
title('Step response for gain of +1 amplifier');grid;figure
pzmap(cloop)
title('Pole map of closed-loop gain of +1 amplifier');grid;figure;
damp(cloop) % Find natural frequency and damping ratio of
H(s)

% gain of +10
f=0.1;
margin(ao*f,d);title('Phase margin calculation for gain of +10
amplifier');figure;
f=tf(f,1);
cloop=feedback(a,f);cloopgainof10=cloop
bode(cloop);title('Bode plot of closed-loop transfer function for gain of
+10 amplifier');figure
step(cloop);grid;title('Step response for gain of +10 amplifier');
pzmap(cloop);title('Pole map of closed-loop gain of +10 amplifier');grid;
damp(cloop)
```

MATLAB script for integral control example

```

function c3
% Control example 3
% driving reactive load

L=10e-6;
C=10e-6;
R=10;
Zo = sqrt(L/C);           % Characteristic Impedance
Q=R/Zo

% Calculate PLANT
num=1;
denom=[L*C L/R 1];
plant=tf(num,denom)
damp(plant)               % Find poles and damping ratio
bode(plant); title('file: c3.m; REACTIVE LOAD EXAMPLE')

% Integral control, attempt #1
Gain=4e3;                 % Integrator gain
denom=[1 0];
Gc=tf(Gain,denom);        % Form Gc(s)
Forw=series(plant,Gc);    % Cascade with plant
margin(Forw);             % Find gain and phase margin
F=tf(1,1);
Cloop=feedback(Forw,F,-1)
figure; step(Cloop);title('STEP RESPONSE, CONTROLLER #1');grid

%Integral control, attempt #2
figure
d=[1/5e4 1];
LPF=tf(1,d);              % Add lowpass filter to damp complex pole pair
Gc=series(LPF,Gc);
Forw=series(plant,Gc);
margin(Forw); title('PHASE MARGIN CALC., CONTROLLER #2');
F=tf(1,1);
Cloop=feedback(Forw,F,-1)
figure; step(Cloop);title('STEP RESPONSE, CONTROLLER #2');grid

```

MATLAB script for MOSFET current source example

```
function moscursource
% Analysis of MOSFET current source
% Marc Thompson, 3/28/00

% LOAD
RL=1;

% MOSFET model
gm = 8.6; % transconductance
Cgs = 1040e-12;
Cgd = 160e-12;

% OPAMP model
rout = 100; % output resistance of opamp
ao=2*pi*4*1e6; % GBP = 4 MHz
denom=[1/ao 0];
highpole=[1/ao 1];
d=conv(denom,highpole);
opamp=tf(1,d)

% MOSFET follower model
Rsense = 0.08; % current sense resistor
Ao=gm*Rsense/(1+gm*Rsense); % gain of follower

% MOSFET OCTC calculation
Rgs=(rout+Rsense)/(1+gm*Rsense);
Tgs=Rgs*Cgs;
GM=gm/(1+gm*Rsense);
Rgd=rout+RL+(GM*rout*RL);
Tgd=Rgd*Cgd;
T=Tgs+Tgd % sum of OCTCs
mosfetpole=1/T
mosfet=tf(Ao,[T 1])

% Find loop transmission
LT=series(opamp,mosfet)
margin(LT);
title('moscursource. FREQUENCY RESPONSE OF UNCOMPENSATED MOSFET CURRENT
SOURCE')
figure;

% close the loop
f=tf(1,1)
uncomp=feedback(LT,f,-1);
step(uncomp); grid;
```

```
title('moscursource. STEP RESPONSE OF UNCOMPENSATED MOSFET CURRENT SOURCE')
figure

% add lag compensation
Rf=470000; C=1e-9;
Ri=47000;
numlag=[Ri*C 1];
denomlag=[(Ri+Rf)*C 1];
f=tf(numlag,denomlag)
comp=feedback(LT,f,-1);
comp=series(f,comp);
step(comp); grid;
title('moscursource. STEP RESPONSE OF LAG COMPENSATED MOSFET CURRENT
SOURCE')
```

MATLAB script for Maglev example

```
function maglev
% Maglev example
% Marc Thompson 4/3/00

% Maglev plant
wn=sqrt(1e3);
num=1.96e-6;
denom=[1/wn^2 0 1];
plant=tf(num,denom);
step(plant);
title('Step response of uncompensated Maglev system');grid;
figure

% feedback compensation
Kp=1e5;
Kv=1e4;
num=[Kv Kp];
f=tf(num,1);

% Closed-loop
sys=feedback(plant,f,-1)
damp(sys)
step(sys);
grid;
title('Step response of compensated Maglev system')
```

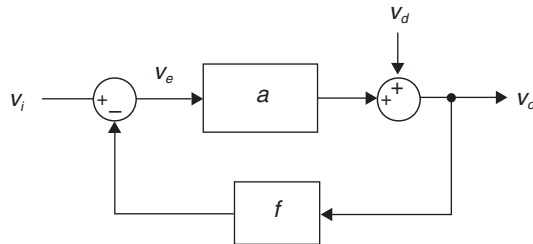

Chapter 11 Problems

Problem 11.1

Consider the negative feedback amplifier in **Figure 11-47**. Assume that the input voltage $v_i = 1\text{V}$ and the disturbance input $v_d = 1\text{V}$. Feedback factor $f = 0.1$.

- (a) For $a = 100$, find output voltage v_o and error voltage v_e .
- (b) For $a = 100,000$ find v_o and v_e .

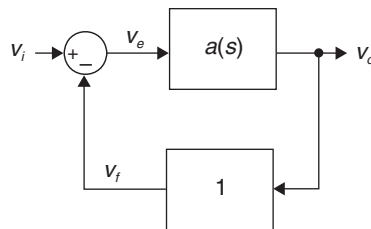
Figure 11-47: Negative feedback amplifier for Problem 11.1.



Problem 11.2

Now, consider the circuit in **Figure 11-48** where $a(s) = K/s$, an integrator with $K = 1000$. There is unity feedback. Find the transfer function $v_o(s)/v_i(s)$. Plot the pole-zero plot and Bode (magnitude and phase) plot of this transfer function.

Figure 11-48: Negative feedback amplifier for Problem 11.2.



Problem 11.3

This problem considers a nonlinear element inside a negative feedback loop (**Figure 11-49**). The incremental gain for the nonlinear element is:

$$v_b/v_a = 0 \text{ for } |v_a| < 0.6\text{V}$$

$$v_b/v_a = 1 \text{ for } 0.6\text{V} < |v_a| < 12\text{V}$$

$$v_b/v_a = 0 \text{ for } |v_a| > 12\text{V}$$

The gain profile of this nonlinear element is typical of a transistor power output stage with crossover distortion.

For this closed-loop amplifier, plot v_o for v_i ranging from -12V to $+12\text{V}$.

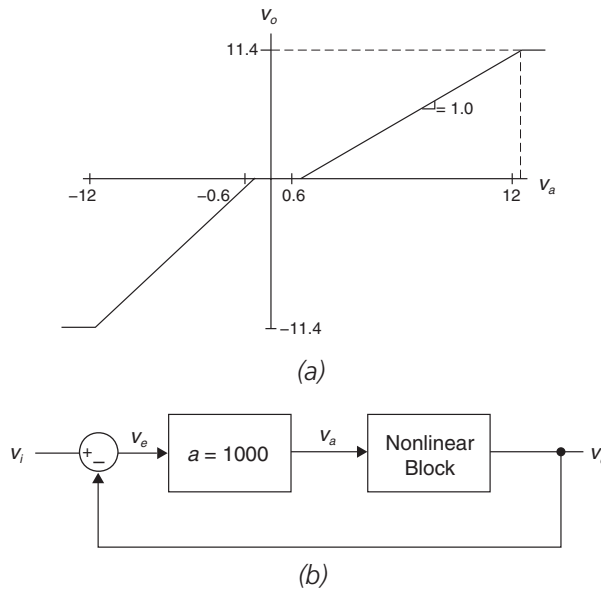


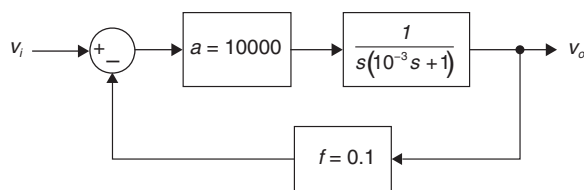
Figure 11-49: Negative feedback amplifier for Problem 11.3.
(a) Nonlinear block. (b) Feedback loop.

Problem 11.4

For the feedback system of **Figure 11-50**:

- Find and plot the ideal input-output transfer function $v_o/v_i(s)$.
- Find the negative of the loop transmission $-L.T.$
- Plot the Bode plot of the loop transmission magnitude and phase.
- Find the crossover frequency and phase margin.

Figure 11-50: Negative feedback amplifier for Problem 11.4.



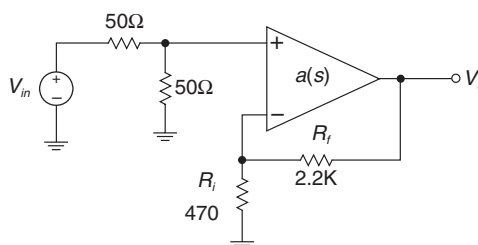
Problem 11.5

The video amplifier shown in **Figure 11-51** is a high-bandwidth, positive-gain stage designed to amplify a video signal v_i . The 50Ω resistive divider models the cabling from the video generator. Assume that the op-amp open-loop transfer function $a(s)$ is:

$$a(s) = \frac{10^5}{(10^{-3}s + 1)(10^{-8}s + 1)}$$

You may assume that all other op-amp parameters are ideal (i.e., infinite slew rate, zero input current, zero output resistance, etc.).

Figure 11-51: Video amplifier for Problem 11.5.



- Draw a block diagram of this system, with the input to your block diagram being signal v_i and the output being v_o .
- Find the ideal closed-loop gain of this configuration. (Ideal meaning that in this part of the problem, you don't consider any bandwidth limitations.)
- Find the negative of the loop transmission ($-L.T.$). Plot $-L.T.$ magnitude and phase and estimate crossover frequency, gain margin and phase margin.
- Now, assume a unit step at the input v_i . Plot $v_o(t)$, assuming that the op-amp does not slew-rate limit.

Problem 11.6

Using the Routh test, find the number of poles in the right-half plane of the transfer function:

$$H(s) = \frac{40}{s^5 + 10s^4 - 5s^3 - 50s^2 + 4s + 40} \quad [11-49]$$

References

- Abramovitch, D., "Phase-locked loops: a control centric tutorial," *Proceedings of the 2002 American Control Conference*, May 8–10, 2002, pp. 1–15.
- , "The outrigger: a prehistoric feedback mechanism," *Proceedings of the 42nd IEEE Conference on Decision and Control*, December 9–12, 2003, pp. 2000–2009.
- Abramovitch, D., and Franklin, G., "A brief history of disk drive control," *IEEE Control Systems Magazine*, vol. 22, no. 3, June 2002, pp. 28–42.
- Bennett, S., "Development of the PID controller," *IEEE Control Systems Magazine*, vol. 13, no. 6, December 1993, pp. 58–62, 64–65.
- Bernstein, D. S., "Feedback Control: An Invisible Thread in the History of Technology," *IEEE Control Systems Magazine*, vol. 22, no. 2, April 2002, pp. 53–68.
- Black, Harold S., "Stabilized Feed-Back Amplifiers," *Electrical Engineering*, vol. 53, no. 1, 1934, pp. 114–120, reprinted in *Proceedings of the IEEE*, vol. 87, no. 2, February 1999, pp. 379–385.
- , "Inventing the Negative Feedback Amplifier," *IEEE Spectrum*, Dec. 1977, pp. 55–60.
- , United States Patent #2,102,671, "Wave Translation System," issued December 21, 1937, available from www.uspto.gov.
- Calleja, Hugo, "An Approach to Amplifier Frequency Compensation," *IEEE Transactions on Education*, vol. 46, no. 1, February 2003, pp. 43–49.
- Denny, Mark, "Watt steam governor stability," *European Journal of Physics*, vol. 23, 2002, pp. 339–351.
- Desoer, C., "In Memoriam: Harold Stephen Black (1898–1983)," *IEEE Transactions on Automatic Control*, vol. 29, no. 8, August 1984, pp. 673–674.
- Fasol, K. H., "A short history of hydropower control," *IEEE Control Systems Magazine*, vol. 22, no. 4, August 2002, pp. 68–76.
- Headrick, M. V., "Origin and evolution of the anchor clock escapement," *IEEE Control Systems Magazine*, vol. 22, no. 2, April 2002, pp. 41–52.
- Herwald, S., "Recollections of the early development of servomechanisms and control systems," *IEEE Control Systems Magazine*, vol. 4, no. 4, November 1984, pp. 29–32.
- Jury, E., "On the history and progress of sampled-data systems," *IEEE Control Systems Magazine*, vol. 7, no. 1, February 1987, pp. 16–21.
- Kline, R., "Harold Black and the Negative-Feedback Amplifier," *IEEE Control Systems Magazine*, vol. 13, no. 4, Aug. 1993, pp. 82–85.
- Lepschy, A. M., Mian, G. A., and Viaro, U., "Feedback control in ancient water and mechanical clocks," *IEEE Transactions on Education*, vol. 35, no. 1, February 1992, pp. 3–10.
- Lewis, F.L., *Applied Optimal Control and Estimation*, Prentice-Hall, 1992.
- Lundberg Kent, "Internal and External Op-Amp Compensation: A Control-Centric Tutorial," *ACC 2004*.

- Lundberg, Kent H., and Roberge, James K.. “Classical Dual-Inverted-Pendulum Control,” *Proceeding of the IEEE CDC 2003*, pp. 4399–4404, December 9–12, 2003, Maui, Hawaii.
- Mancini, Ron, “The Saga of Harry Black,” *EDN Magazine*, March 15, 2001, p. 34.
- Maxwell, James C., “On Governors,” *Proceedings of the Royal Society*, 1867, pp. 270–283.
- Mayr, Otto, *The Origins of Feedback Control*, The MIT Press, 1970.
- Michel, A. N., “Stability: the common thread in the evolution of feedback control,” *IEEE Control Systems Magazine*, vol. 16, no. 3, June 1996, pp. 50–60.
- Pidhayny, D., “The origins of feedback control,” *IEEE Transactions on Automatic Control*, vol. 17, no. 2, April 1972, pp. 283–284.
- Roberge, James K., *Operational Amplifiers: Theory and Practice*, John Wiley, 1975.
- Siebert, William McC., *Circuits, Signals and Systems*, The MIT Press, 1986.
- Tilbury, Dawn, Luntz, Jonathan, and Messner, William, “Controls Education on the WWW: Tutorials for MATLAB and Simulink,” *Proceedings of the American Control Conference*, Philadelphia PA, June 1998, pp. 1304–1308.

Basic Operational Amplifier Topologies and a Case Study

In This Chapter

- The basic operational amplifier is discussed from a topological point-of-view. A step-by-step case study illustrates the basic building blocks in a monolithic op-amp. At the end of this chapter we'll consider some of the real-world limitations of operational amplifiers.

Basic Device Operation

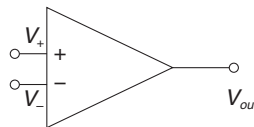
The ideal operational (**Figure 12-1**) amplifier has the following characteristics:

- *Differential inputs.* The output is an amplified version of the difference between the + and – terminals.
- *Infinite gain.* The gain is infinite.
- *Infinite bandwidth.* There are no bandwidth limitations.
- *Infinite slew rate.* There is no limit to the rate with which the output can change. In other words, there is no limit to dV_{out}/dt .
- *Zero input current.* The input current to both inputs is zero.
- *Zero output impedance.* The output impedance is zero.
- *Zero power dissipation.* The ideal op-amp doesn't draw or dissipate any power.
- *Infinite power supply rejection.* The output is not dependent on variations in power supply voltage.
- *Infinite common-mode signal rejection.* The output doesn't depend on the value of the common-mode signal.

The ideal op-amp is, of course, nonexistent, but device manufacturers have done a better and better job over the years designing devices that approach the ideal. For instance, it is common to find devices with DC gains much higher than 10^6 and/or gain/bandwidth products of greater than 100 MHz.¹

¹ Compare this to the baseline specs of the 741 op-amp (a device from the 60s still for sale), with nominal DC gain of 200,000 and gain-bandwidth product of 1 MHz. A Linear Technology LT1226, (a 1990s-era device) has a gain-bandwidth product of 1000 MHz.

Figure 12-1: Ideal operational amplifier showing differential inputs V_+ and V_- . The ideal op-amp has zero input current and a gain approaching infinity that amplifies the difference between V_+ and V_- .



The usual method for doing first-cut analysis of closed-loop op-amp circuits is to assume a “virtual ground.” This term is a bit of a misnomer, because the input terminals in general don’t need to be at ground potential. However, in an operational amplifier, operating with negative feedback, the difference between the two inputs is ideally zero volts. If the op-amp + terminal is at ground, the – terminal will be at approximately ground. If, in a different configuration the + terminal is at +6V, the – terminal will also be at approximately +6V.

A basic two-stage op-amp is shown in **Figure 12-2**. This is a two-stage op-amp because it consists of two gain stages: the input differential gain stage followed by a second common-emitter gain stage. The input differential amplifier stage (Q_1 and Q_2) has high differential-mode gain and low common-mode gain. The second gain stage (Q_3) provides additional gain and also provides a DC level shifting function. Compensation capacitor C_c provides a low-frequency pole and by the process of pole-splitting² causes the next-highest frequency pole to move to a higher frequency. Pole splitting has important ramifications in overall amplifier stability.

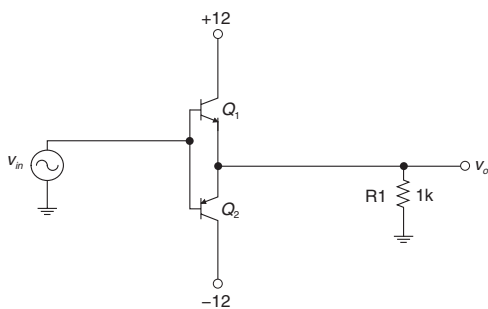
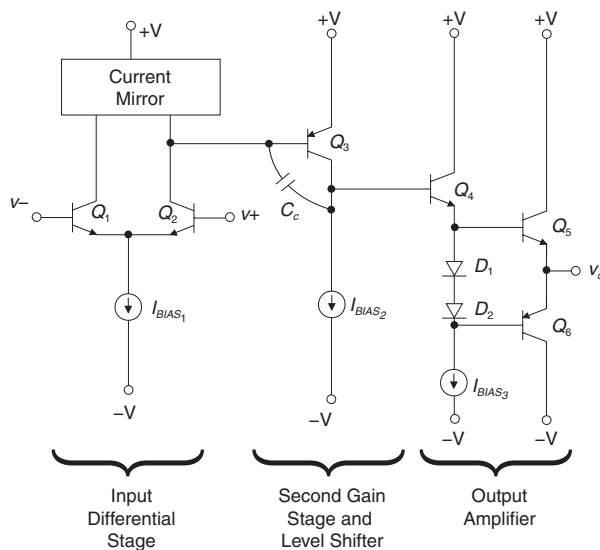
Emitter follower Q_4 buffers the high gain node from the output. The output stage (Q_5 and Q_6) is a class AB *push-pull* stage. This output stage can source or sink current. For instance, when sourcing current Q_5 is ON; when sinking current Q_6 is ON. The diodes at the bases of Q_5 and Q_6 provide two functions. First, they reduce *crossover distortion* in the output stage. Secondly, by proper sizing D_1 and D_2 ³ relative to the sizes of Q_5 and Q_6 , we can set a modest bias current in the output stage transistors, which lowers the incremental output resistance of the output amplifier.

A push-pull output stage without biasing diodes is shown in the PSPICE circuit of **Figure 12-3**. The output waveforms show significant DC offset as well as *crossover distortion* or a dead zone of approximately $\pm 0.6\text{V}$ around $v_o = 0$.

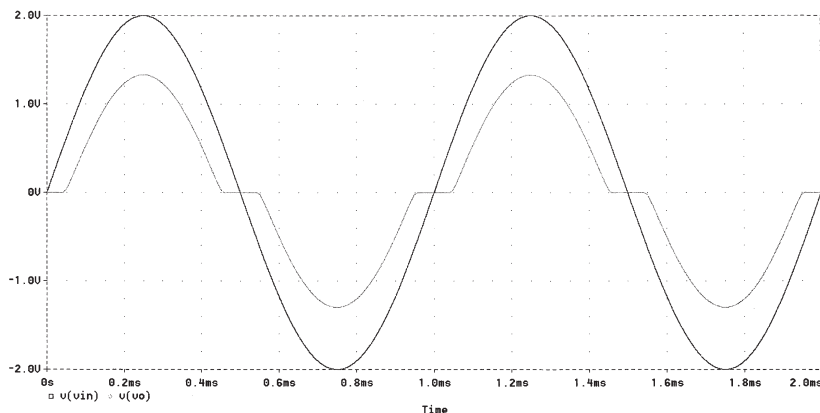
² We discussed pole splitting in detail in Chapter 7.

³ In some real-world op-amps, the voltage drop provided by these two diodes is alternately provided by an alternate circuit topology. See, e.g., the 741 operational amplifier, which uses a “ V_{BE} multiplier” to provide this function.

Figure 12-2: Basic two-stage op-amp with an input differential gain stage followed by a second gain and level shifting stage and an output buffer.



(a)



(b)

Figure 12-3: Push-pull stage. (a) Basic circuit. (b) Output response for 4V pp sinewave input exhibiting crossover distortion at the output.

This push-pull circuit can be modified to improve voltage offset as in **Figure 12-4**. The diodes between the bases of the transistors set up an approximate 1.2V bias that begins to turn on the output transistors. Resistors R_1 and R_2 bias the diode string. The small resistor R_3 is provided to reduce the chances of thermal runaway. Here's how thermal runaway would work in this circuit: Let's say that the output transistors Q_3 and Q_4 are carrying significant current. They will heat up. Since the transistors are now at a higher temperature, they require less base-emitter voltage to sustain the same output current.⁴ Therefore, the current in the output stage increases. The transistors heat up more, and the circuit “runs-away” thermally. The inclusion of the small resistor R_3 ensures that if the collector currents begin to increase significantly that the voltage drop across R_3 will become significant, hence stealing away base drive voltage from the transistors.

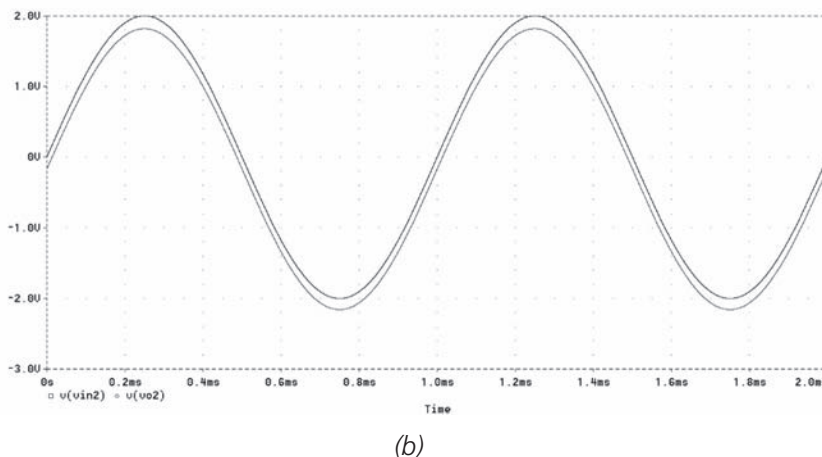
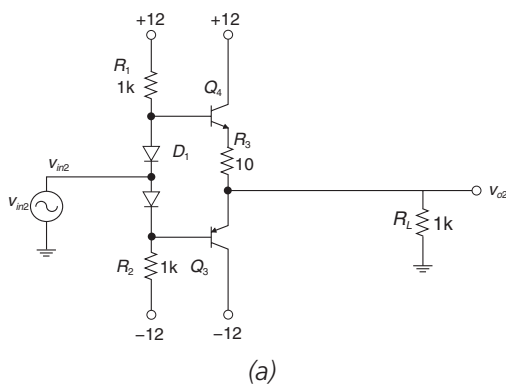


Figure 12-4: Push-pull stage with base biasing diodes.
(a) Basic circuit. (b) Output response for 4V pp sinewave input

⁴ Remember that the temperature coefficient of V_{BE} is approximately $-2.2 \text{ mV/}^\circ\text{C}$.

Let's next study the power dissipation in the basic push-pull amplifier (**Figure 12-5a**), assuming ideal components (i.e., no crossover distortion) and a symmetric power supply voltage. The NPN transistor is on for positive load current I_L . The power dissipation is:

$$P_D = V_{CE} I_L = (V_S - V_o) I_L = V_S I_L - I_L^2 R_L \quad [12-1]$$

Therefore, the shape of the power dissipation vs. load current curve has a parabolic shape (**Figure 12-5b**). The maximum power dissipation is found by taking the derivative of the power dissipation with respect to load current.

$$\frac{dP_D}{dI_L} = V_S - 2I_L R_L \quad [12-2]$$

The maximum power dissipation occurs at load current $I_L = V_S/(2R_L)$, and at an output voltage which is half the supply voltage.⁵

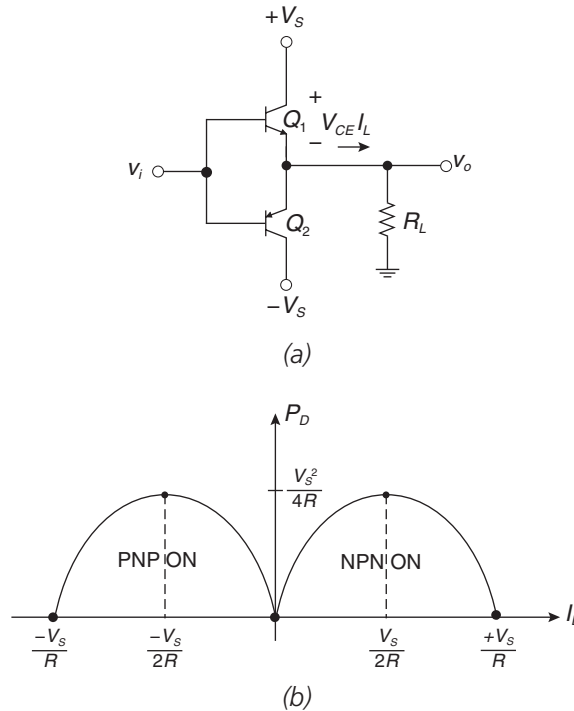


Figure 12-5: Circuit for studying power dissipation in push-pull amplifier. (a) Basic circuit. (b) Transistor power dissipation P_D vs. load current I_L for NPN and PNP transistors.

⁵ One might ask why the power dissipation is zero at maximum collector current. In this simplified model, we assume that the saturation voltage of the transistor is zero; hence a finite collector current multiplied by zero V_{CE} results in zero transistor power dissipation.

Example 12.1: Case study: Design, analysis and simulation of a discrete operational amplifier

Let's start off designing a discrete operational amplifier using discrete components. This resulting design will not be state-of-the-art, but rather the design process illustrates the various building blocks that exist in most monolithic IC op-amps. The design insight afforded by this exercise will help us to understand the limitations of real-world op-amps. Op-amp limitations are discussed in more detail later on in this chapter.

Differential input stage

In an op-amp, some form of differential input stage is needed. One possible implementation is the current-mirror-loaded differential input stage (**Figure 12-6**). The components of this stage are:

- Differential input transistors Q_1 and Q_2
- Current mirror transistors Q_3 and Q_4
- Bias current source I_{bias} , with details omitted here.

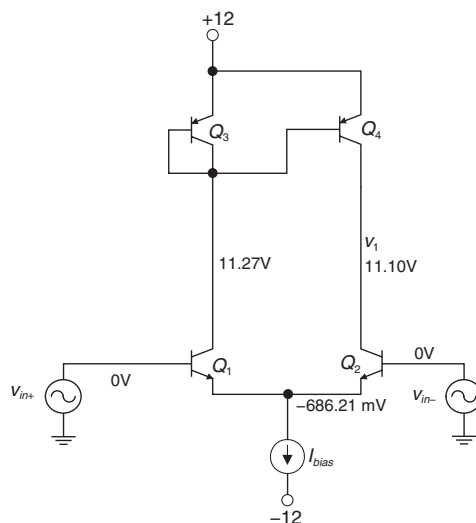


Figure 12-6: Differential input stage with input transistors Q_1 and Q_2 , and current mirror Q_3 and Q_4 . We take the output voltage to be v_1 , at the connection of the collectors of Q_2 and Q_4 .

Let's do a DC sweep of v_{in+} while grounding v_{in-} . This will give us an idea as to the voltage offset of the input stage. Ideally, we would like the output v_1 to be in the middle of the (approximately) linear gain region $v_{in+} = 0V$. In practicality, however, this voltage offset will not be zero due to device mismatches and errors in the Q_3/Q_4 current mirror. A SPICE simulation shows this effect; the output at v_1 transitions from approximately ground to approximately +12V while v_{in+} is varied from approximately -15 mV to ground. Note that the gain of this stage (gain being v_{o1}/v_{in+}) is approximately 1000.

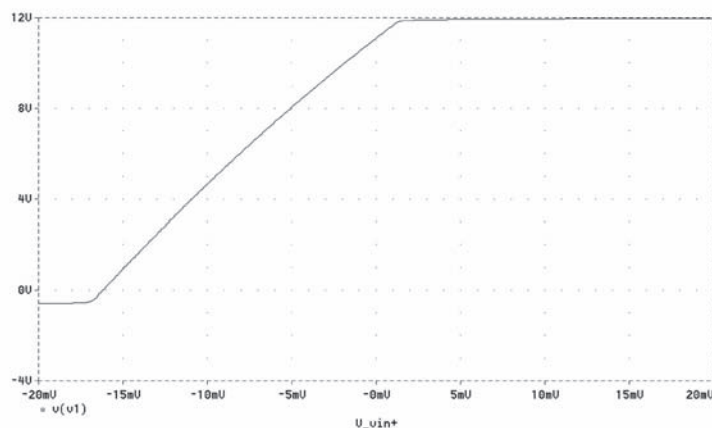
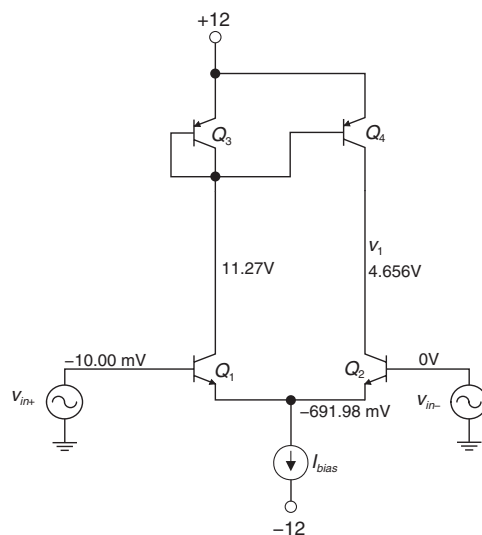
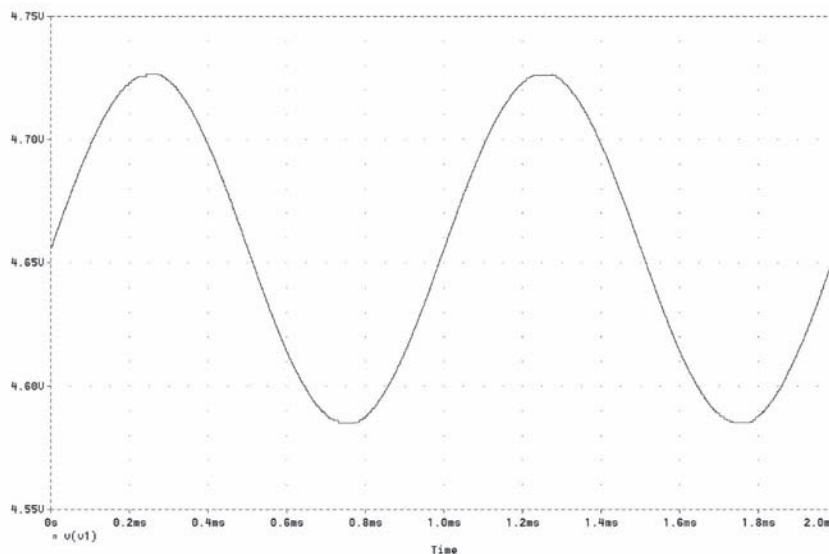


Figure 12-7: DC Sweep of differential input stage.

Let's put a DC bias of -10 mV at v_{in+} , and also put a 200- μV peak-peak 1-kHz sinewave at the input (**Figure 12-8a**). Note that the DC bias level of v_1 sits at +4.656V (**Figure 12-8b**), which means that Q_2 and Q_4 are both ON and in the forward-active region. Therefore, we expect active gain with the transistors biased under these conditions. The gain from v_{in+} to v_1 is approximately 575.



(a)



(b)

Figure 12-8: Results of AC sinewave sweep. (a) Circuit.
(b) PSPICE result for a 0.2 mV-pp variation in v_{in+} .

Continuing on with the design, we now recognize that further transistors are needed for buffering and level-shifting functions. The output at v_1 only has positive voltages; in a practical amplifier, we transition both positive and negative voltages at the output. Therefore, we need to level-shift the output voltage from the differential stage. One possible way to do this

is with a folded-cascode amplifier, as shown in **Figure 12-9**. This folded cascode (Q_5 and bias current sources I_{bias2} and I_{bias3}) works as follows: First, assume that I_{bias2} and I_{bias3} are the same value. Transistor Q_5 buffers the difference between the collector currents of Q_4 and Q_5 , or in this case the small-signal variation in i_{c5} is:

$$i_{c5} \approx I_{C4} - I_{C2} = \left(\frac{I_{BIAS}}{2} + \Delta i \right) - \left(\frac{I_{BIAS}}{2} - \Delta i \right) = 2\Delta i \quad [12-3]$$

Furthermore, for the differential input stage:

$$\Delta i = g_m v_{in} \quad [12-4]$$

It is the $2\Delta i$ current that is converted to voltage v_2 by loading it with current source I_{BIAS3} .

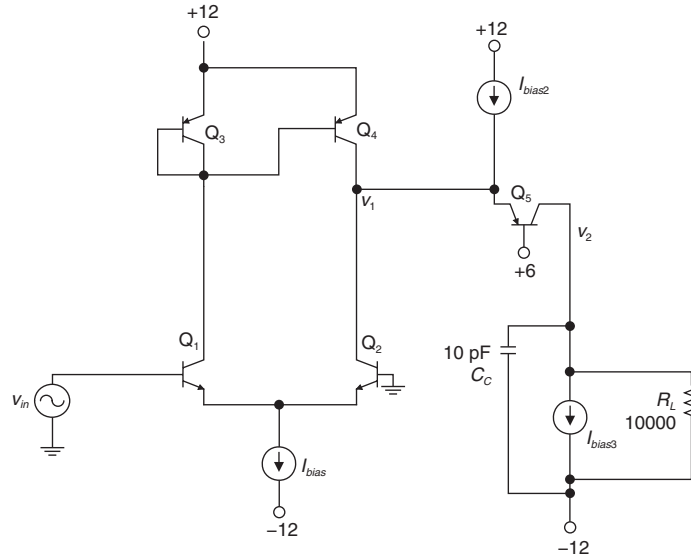
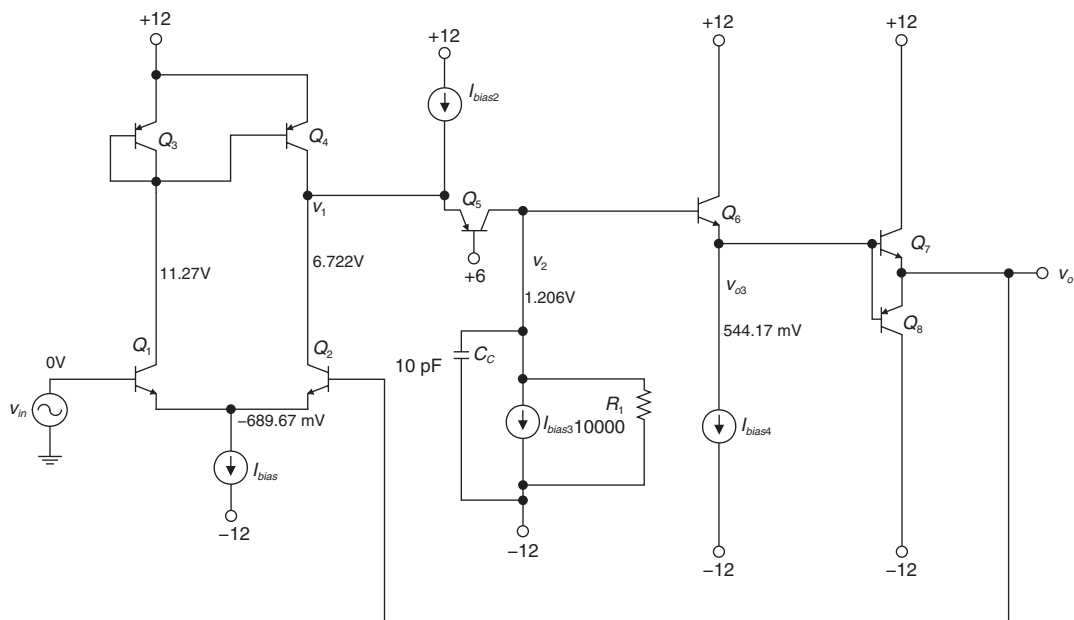


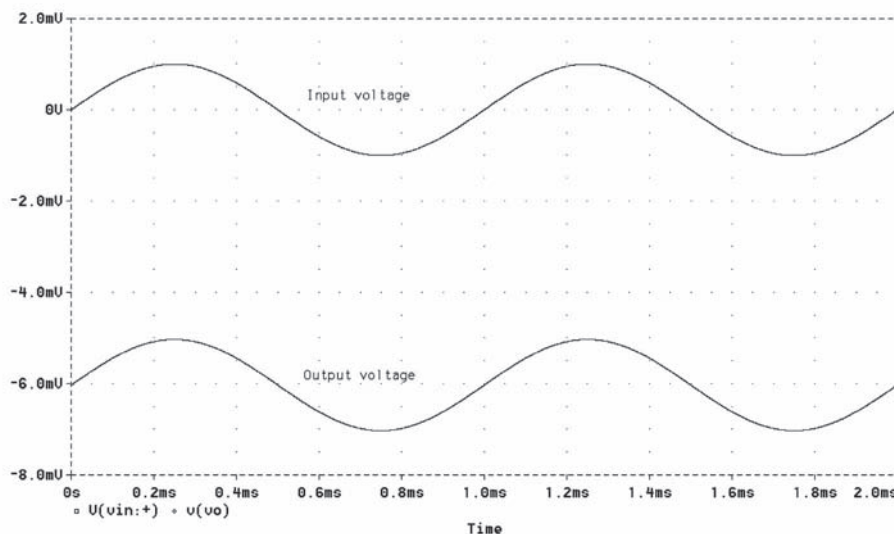
Figure 12-9: Differential amplifier plus folded cascode

Emitter follower buffering and output push-pull stage

To this input stage, we now need to add an output push-pull stage (**Figure 12-10a**). The output transistors Q_7/Q_8 can sink and source current. Emitter follower Q_6 is added so that the output transistor doesn't significantly load down the high-gain node at the collector of Q_5 . We see the input and output of this closed-loop gain-of-1 configuration in **Figure 12-10b**. We note that there is a -6 mV offset between input and output.



(a)

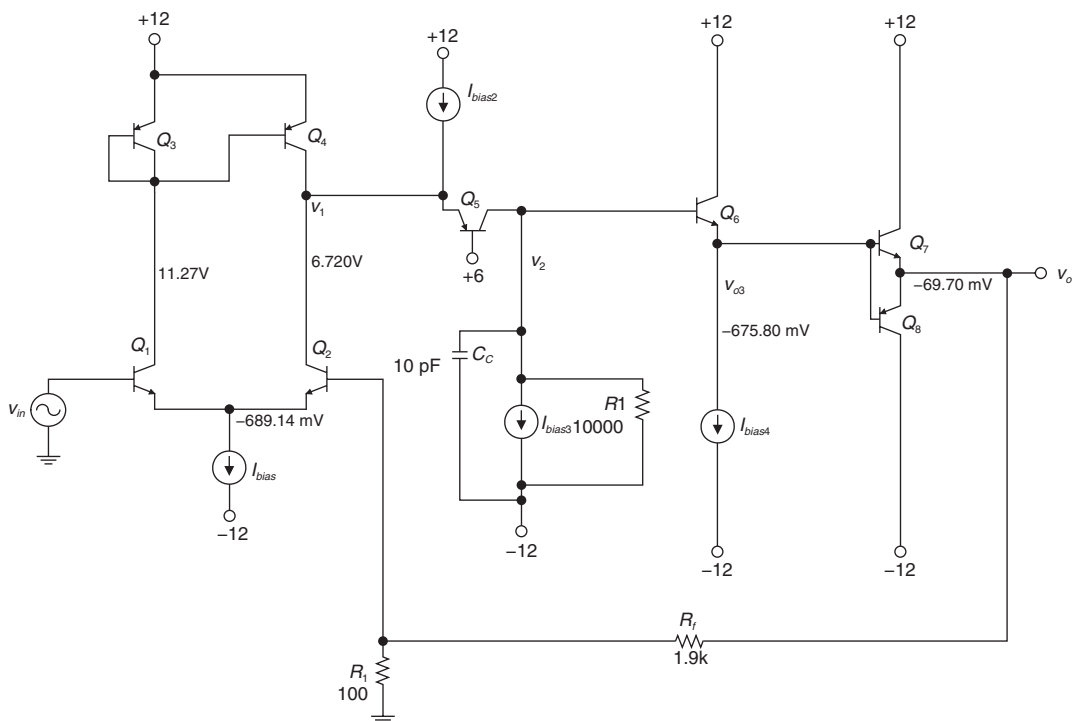


(b)

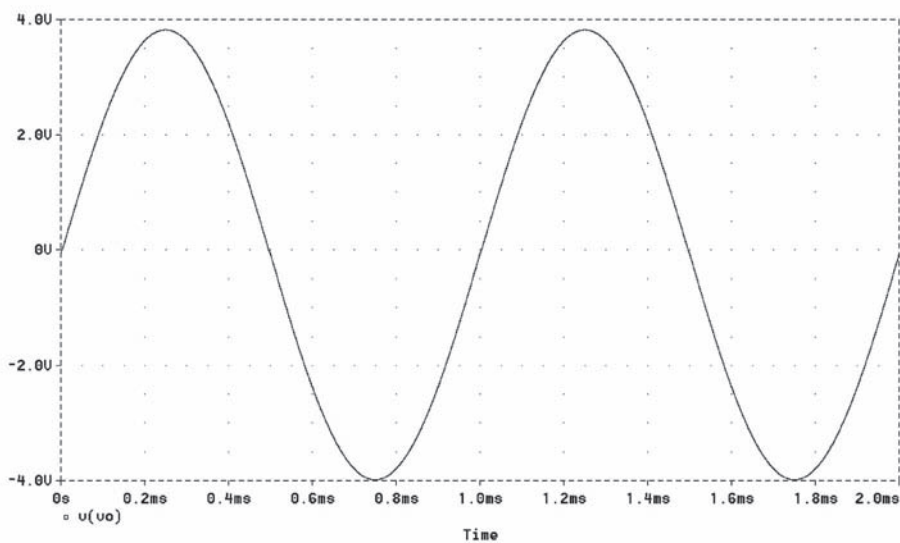
Figure 12-10: Final amplifier showing unity feedback.

(a) Circuit showing some PSpice-predicted node voltages. (b) Input and output voltages.

We next configure the op-amp for a gain-of-20 (**Figure 12-11a**). We note that the output can swing $\pm 8V$ without clipping (**Figure 12-11b**). As we raise the amplitude of the driving signal further, the output clips at $\sim 5V$ (**Figure 12-11c**) but the negative sinewave is OK.



(a)

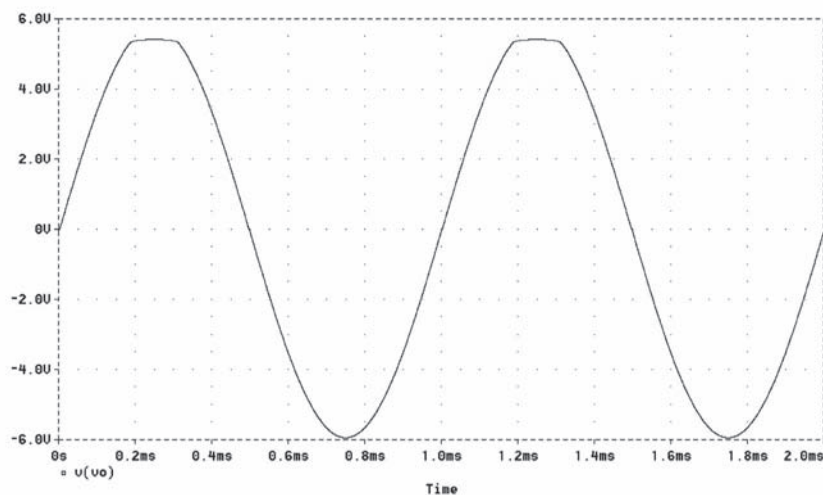


(b)

Figure 12-11: Final amplifier in a gain-of-20 configuration.

(a) Circuit showing some PSPICE-predicted node voltages.

(b) Output for input 400-mV pp sinewave, 1 kHz. (Continued on following page.)



(c)

Figure 12-11 (continued): (c) Output for input 600-mV pp sinewave.

Note: This design is shown for illustrative purposes only. There are many design improvements that could be made, including better biasing, more intelligent topology selection, and the like. So, please, no diatribes about how it's a crummy design!

Brief Review of LM741 Op-Amp Schematic

We'll now do a quick review of the topology of the LM741 operational amplifier, made by many companies since the 1960s (**Figure 12-12**). This is a two-stage op-amp with pole splitting.

Input differential gain stage: Q_1 , Q_2 , Q_3 and Q_4 . Q_5 and Q_6 is the current mirror load for the differential stage. Q_7 is a beta helper transistor that improves the gain ratio of the mirror. Q_8 is the bias current source.

Bias current mirrors: Q_8 and Q_9 ; Q_{10} and Q_{11} ; Q_{12} and Q_{13} . These mirrors set the bias current levels throughout the op-amp.

Second gain stage: Comprised of common emitter amplifier with 50Ω emitter degeneration Q_{17} . Transistor Q_{15} provides buffering between the output of the first gain stage and the input of the second gain stage. The 30-picofarad capacitor C_1 provides a pole-splitting function.

Output push-pull amplifier. Transistors Q_{14} and Q_{20} form the heart of the push-pull. A V_{BE} multiplier⁶ between the bases of Q_{14} and Q_{20} partially biases the push-pull ON. Transistor Q_{15} provides current limiting for positive output currents.

⁶ This " V_{BE} multiplier" circuit provides approximately $1.6 V_{BE}$ drop.

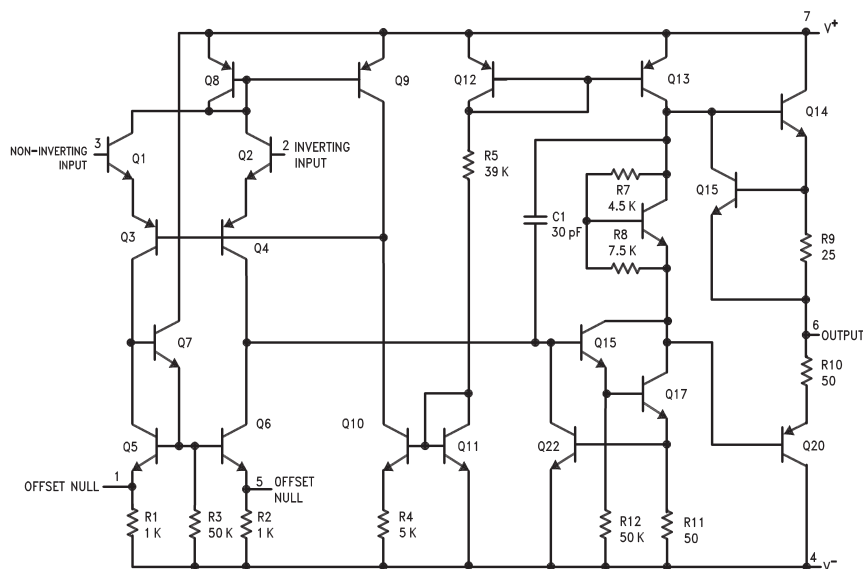


Figure 12-12: Schematic of the LM741 op-amp, from National Semiconductor. Reprinted with permission of National Semiconductor Corporation.

Some Real-World Limitations of Operational Amplifiers

Voltage offset

In Chapter 5, we analyzed the differential amplifier, and assumed ideal (i.e., perfectly matched) devices. However, in the real world, there is mismatch between devices, and in the operational amplifier front-end, this mismatch manifests itself as a voltage offset. The voltage offset is the voltage that must be applied differentially to force the output of the op-amp to zero. In commercially available op-amps, the voltage offset is typically fractions of a millivolt, up to a few millivolts.

We can model voltage offset by adding a voltage generator of value V_{os} to an ideal op-amp, as shown in **Figure 12-13**. The voltage offset can be especially troublesome in high-gain configurations, such as **Figure 12-14**. Note that the offset generator is in series with the input signal v_i and that output voltage of this configuration is:

$$v_o = 1000v_i + 1000V_{os} \quad [12-5]$$

For small input signals, the voltage offset can swamp out the input signal.

Figure 12-13: Operational amplifier showing voltage source modeling the voltage offset V_{os} .

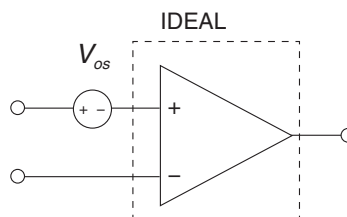
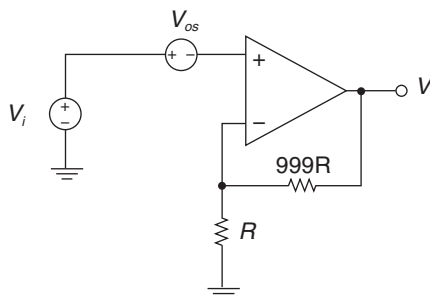


Figure 12-14: Gain of +1000 amplifier showing effects of voltage offset V_{os} .



A datasheet excerpt from National Semiconductor shows the input offset voltage for the LM741 op-amp (**Figure 12-15**). Note that different grades of op-amp have slightly different specifications, but the typical offset is ~1 millivolt while the maximum offset is a few millivolts.

Parameter	Conditions	LM741A			LM741			LM741C			Units
		Min	Typ	Max	Min	Typ	Max	Min	Typ	Max	
Input Offset Voltage	$T_A = 25^\circ\text{C}$										
	$R_S \leq 10\text{ k}\Omega$					1.0	5.0		2.0	6.0	mV
	$R_S \leq 50\Omega$		0.8	3.0							mV
	$T_{AMIN} \leq T_A \leq T_{AMAX}$										
	$R_S \leq 50\Omega$			4.0							mV
	$R_S \leq 10\text{ k}\Omega$						6.0			7.5	mV

Figure 12-15: Datasheet excerpt from the LM741 operational amplifier (from National Semiconductor) showing voltage offset.
(Reprinted with permission of National Semiconductor Corporation.)

Voltage offset drift with temperature

Another important design issue is voltage offset drift with temperature. Referring to

Figure 12-16, we see a differential input stage typical of that in op-amps. We note that the voltage offset is:

$$V_{os} = V_{BE1} - V_{BE2} \quad [12-6]$$

This voltage is the differential voltage that occurs when the output differential voltage is zero. For a transistor under forward bias there is the exponential relationship between collector current and base-emitter voltage, which for transistor Q_1 is:

$$I_{C1} = I_S \left(e^{\frac{qV_{BE1}}{kT}} - 1 \right) \approx I_S e^{\frac{qV_{BE1}}{kT}} \quad [12-7]$$

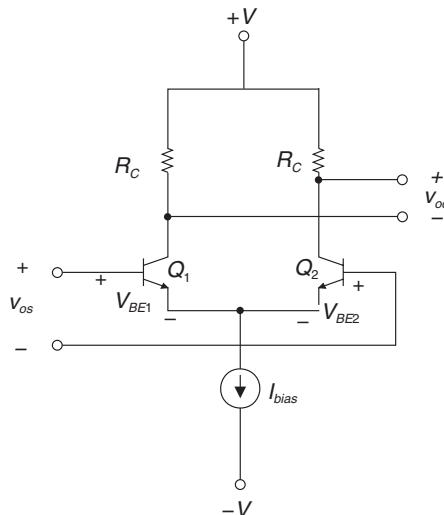
where I_S is the reverse saturation current of the transistor. We now find the equation for the base-emitter voltage of bipolar transistor Q_1 :

$$V_{BE1} = \frac{kT}{q} \ln \left(\frac{I_{C1}}{I_S} \right) \quad [12-8]$$

The temperature coefficient of a single V_{BE} is found by:

$$\frac{dV_{BE1}}{dT} = \frac{k}{q} \ln \left(\frac{I_{C1}}{I_S} \right) - \frac{kT}{qI_S} \frac{dI_S}{dT} = \frac{V_{BE1}}{T} - \frac{kT}{qI_S} \frac{dI_S}{dT} \quad [12-9]$$

Figure 12-16: Circuit for finding offset voltage drift with temperature in a bipolar op-amp.



The second term in this expression is the leakage current temperature coefficient. We can now find the total voltage offset drift of the differential amplifier as:

$$\frac{dV_{os}}{dT} \approx \frac{dV_{BE1}}{dT} - \frac{dV_{BE2}}{dT} = \left(\frac{V_{BE1} - V_{BE2}}{T} \right) - \frac{kT}{qI_{S1}} + \frac{kT}{qI_{S2}} \quad [12-10]$$

In matched transistors, $I_{S1} = I_{S2}$ and the component of voltage offset drift due to leakage current drift cancels, leaving us with:

$$\frac{dV_{os}}{dT} \approx \frac{V_{BE1} - V_{BE2}}{T} \approx \frac{V_{os}}{T} \quad [12-11]$$

Using this expression, let's predict the voltage offset drift for a 741 op-amp with a maximum 4-mV offset at room temperature:

$$\frac{dV_{os}}{dT} \approx \frac{4 \text{ mV}}{300 \text{ K}} \approx 13 \frac{\mu\text{V}}{\text{K}} \quad [12-12]$$

This value is consistent with the datasheet value from the 741 datasheet (**Figure 12-17**).

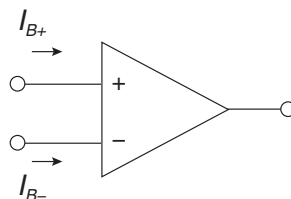
Average Input Offset Voltage Drift				15						$\mu\text{V}/^\circ\text{C}$
------------------------------------	--	--	--	----	--	--	--	--	--	------------------------------

Figure 12-17: Datasheet excerpt from the LM741 operational amplifier (from National Semiconductor) showing voltage offset drift with temperature.
(Reprinted with permission of National Semiconductor Corporation.)

Input bias and input offset current

The input differential amplifier of an op-amp requires base current (in the case of a bipolar-input op-amp). This input current is specified on an op-amp datasheet as *input bias* current, as shown in **Figure 12-18**. The bias currents are small but finite. Datasheets also specify *input offset* current, which is the guaranteed maximum difference between I_{B+} and I_{B-} .

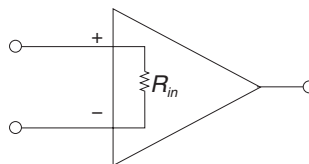
Figure 12-18: Operational amplifier showing small but finite input bias currents I_{B+} and I_{B-} .



Differential input resistance

There is a differential-mode input resistance that can be modeled as a large-valued resistance across the op-amp input terminals.

Figure 12-19: Operational amplifier showing differential input resistance.



Slew rate

We have seen before that a small capacitor is needed to tailor the frequency response of the operational-amplifier open-loop characteristic. By pole splitting, we create a dominant low-frequency pole. Another effect of this feedback capacitance is slew rate limiting. There is a finite amount of current available from the input stage to supply this feedback capacitance. Hence, the output voltage time rate of change is limited by the capacitance and the current available to charge this capacitance, as:

$$\frac{dv}{dt} = \frac{I}{C} \quad [12-13]$$

Slew rate limiting manifests itself when we attempt to switch large amplitude signals. Consider the voltage follower of **Figure 12-20**, where we are switching a 0 to 10V signal. We note that the output shows slew-rate limiting on the positive-going and negative-going edges. Note that the slew rates in the two directions aren't necessarily the same.

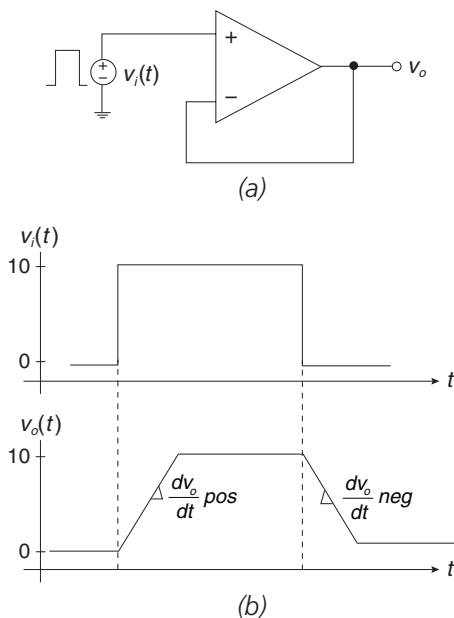
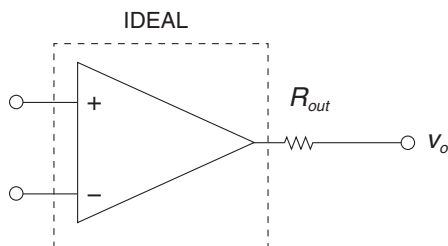


Figure 12-20: Circuit that shows effects of slew rate limiting.

Output resistance and capacitive loading

All operational amplifiers have a finite output resistance. As we'll see, this finite output resistance can have a significant effect on closed-loop stability when driving capacitive loads.

Figure 12-21: Operational amplifier showing output resistance.



Example 12.2: Op-amp driving capacitive load

Let's consider a typical operational amplifier open-loop transfer function:

$$a(s) \approx \frac{10^5}{(0.01s + 1)(10^{-7}s + 1)} \quad [12-14]$$

The open-loop transfer function $a(s)$ of this op-amp is plotted in **Figure 12-22**, where we see the low-frequency pole at 100 radians/second and high-frequency pole at 10^7 radians/second. This type of transfer function is typical of many commercially available op-amps.

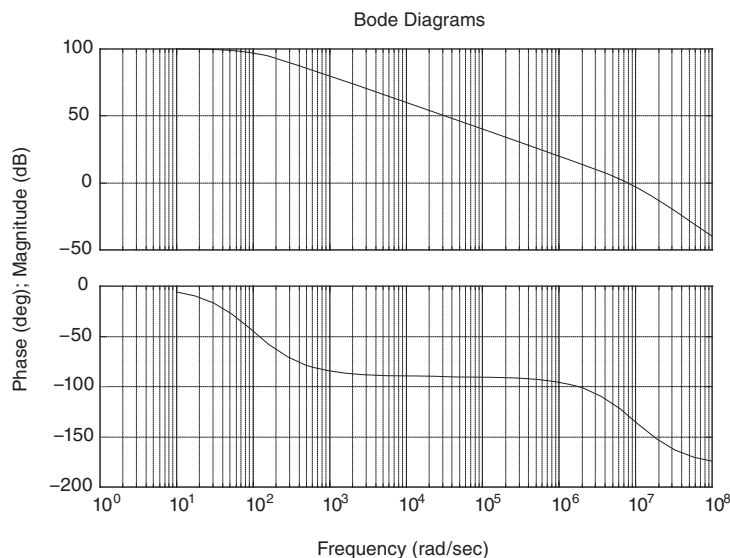


Figure 12-22: Plot of $a(s)$ for typical op-amp.

We'll now see what happens if we use this op-amp to drive a capacitive load in a follower configuration (**Figure 12-23a**). A model of the op-amp showing the op-amp output resistance is shown in **Figure 12-23b**. In this example we'll assume that the output resistance of the op-amp is 100Ω .

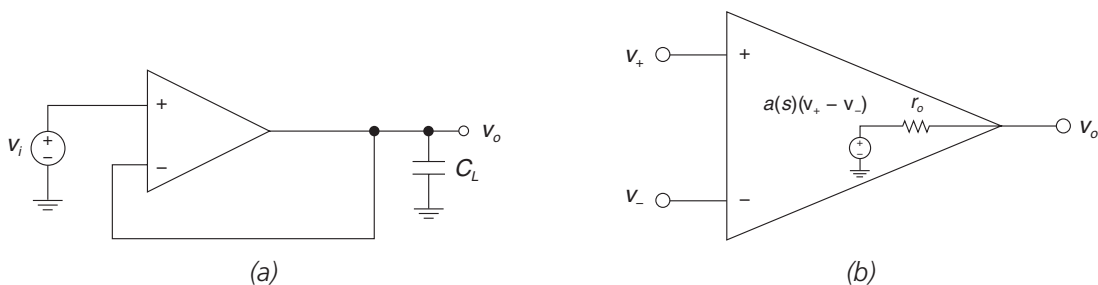


Figure 12-23: Op-amp follower driving capacitive load.
(a) Circuit. (b) Model showing op-amp output resistance r_o .

The block diagram of this capacitively loaded amplifier is shown in **Figure 12-24**.

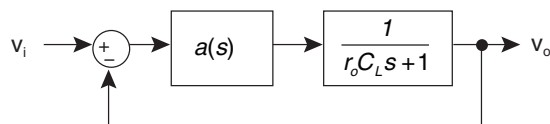


Figure 12-24: Capacitively loaded op-amp follower block diagram.

When we plot the magnitude and phase of the loop transmission (**Figure 12-25**) we note that there is poor phase margin of 22° . The resultant step response of the closed-loop system (**Figure 12-26**) shows significant overshoot, as expected, with this low value of phase margin. The low phase margin is due to the low-pass filter inside the feedback loop, due to the op-amp output resistance interacting with the load capacitor.

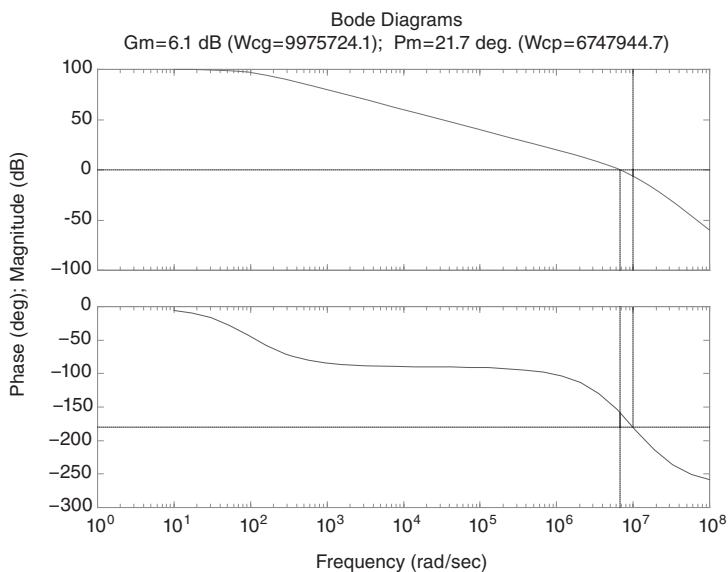


Figure 12-25: Plot of loop transmission gain and phase, showing phase margin of op-amp when loaded with 1000 pF of 21.7° .

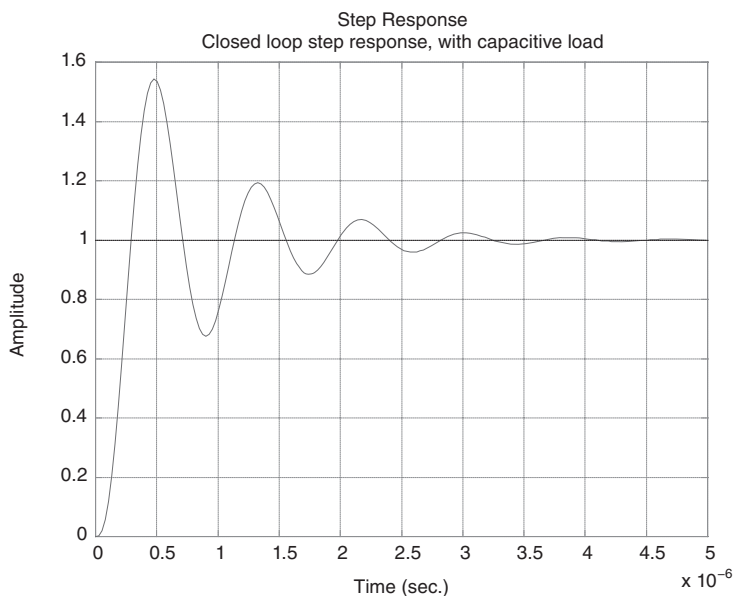


Figure 12-26: Unit step response of capacitively loaded follower.

Chapter 12 Problems

Problem 12.1

Referring to the $\mu\text{A}741\text{A}$ op-amp datasheet, find the maximum frequency (in Hz) that you can drive a 20V pp sinewave without slew rate limiting, using this op-amp. Use typical numbers from the datasheet to determine this number.

Problem 12.2

The 741A circuit of **Figure 12-27** drives a resistive load. The input is a 20V pp sinewave at 100 Hz. Again, using typical numbers from the datasheet, plot the output voltage as a function of time

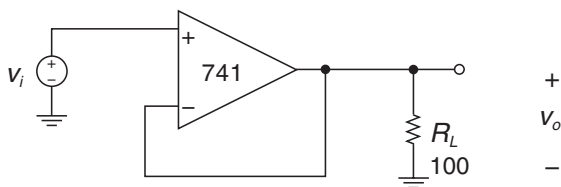


Figure 12-27: Op-amp driving resistive load.

Problem 12.3

In the 741A circuit of **Figure 12-28** find the output voltage. Include the effects of op-amp input bias current as well as the voltage offset of the op-amp. Use typical values from the datasheet.

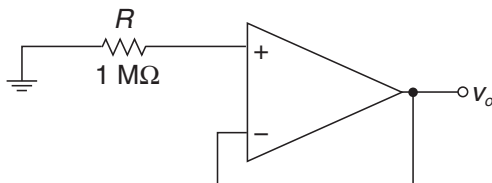


Figure 12-28: Circuit illustrating effects of input bias current and voltage offset of op-amps.

Problem 12.4

This problem revisits this issue of an op-amp follower driving a capacitive load. One way to mitigate the destabilizing effects of driving a capacitive load is to isolate the capacitive load C_L from the op-amp follower circuit using an external resistor R_{EXT} , as in **Figure 12-29**. Draw the block diagram including signals v_i , v_o and v_f , and comment on how this scheme can be used to stabilize this circuit by improving its phase margin.

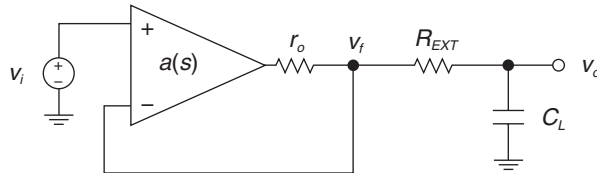


Figure 12-29: Mitigating strategy to improve stability of op-amp follower driving capacitive load C_L . The op-amp has output resistance r_o intrinsic to the device and has transfer function $a(s)$.

References

A wealth of information is provided in the following references. Some of the author's favorites are the references by Bob Widlar, one of the original architects of operational amplifiers. The Solomon and Gray and Meyer references are also excellent overviews of op-amp technology.

- Allen, P. E., "Slew Induced Distortion in Operational Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 12, no. 1, February 1977, pp. 39–44.
- Bowers, D. F., and Wurcer, S. A., "Recent Developments in Bipolar Operational Amplifiers," *Proceedings of the 1999 Bipolar/BiCMOS Circuits and Technology Meeting*, September 26–28, 1999, pp. 38–45.
- Brown, J. L., "Differential Amplifiers that Reject Common-Mode Currents," *IEEE Journal of Solid-State Circuits*, vol. 6, no. 6, December 1971, pp. 385–391.
- Chuang, C. T., "Analysis of the Settling Behavior of an Operational Amplifier," *IEEE Journal of Solid-State Circuits*, vol. 17, no. 1, February 1982, pp. 74–80.

- Comer, D. T., and Comer, D. J., "A New Amplifier Circuit with Both Practical and Tutorial Value," *IEEE Transactions on Education*, vol. 43, no. 1, February 2000, p. 25.
- Erdi, G., "Common-Mode Rejection of Monolithic Operational Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 5, no. 6, December 1970, pp. 365–367.
- Gray, P., and Meyer, R., "Recent Advances in Monolithic Operational Amplifier Design," *IEEE Transactions on Circuits and Systems*, vol. 21, no. 3, May 1974, pp. 317–327.
- Hearn, W. E., "Fast Slewing Monolithic Operational Amplifier," *IEEE Journal of Solid-State Circuits*, vol. 6, no. 1, February 1971, pp. 20–24.
- Huijsing, J. H., and Tol, F., "Monolithic Operational Amplifier Design with Improved HF Behaviour," *IEEE Journal of Solid-State Circuits*, vol. 11, no. 2, April 1976, pp. 323–328.
- Ruediger, V. G., Hosticka, V. G. and B. J., "The Response of 741 Op Amps to Very Short Pulses," *IEEE Journal of Solid-State Circuits*, vol. 15, no. 5, October 1980, pp. 908–910.
- Solomon, J. E., "The Monolithic Op Amp: a Tutorial Study," *IEEE Journal of Solid-State Circuits*, vol. 9, no. 6, December 1974, pp. 314–332.
- Soloman, James E., "A Tribute to Bob Widlar," *IEEE Journal of Solid State Circuits*, vol. 26, no. 8, August 1991, pp. 1087–1089.
- Soundararajan, K., and Ramakrishna, K., "Characteristics of Nonideal Operational Amplifiers," *IEEE Transactions on Circuits and Systems*, vol. 21, no. 1, January 1974, pp. 69–75.
- Treleaven, D., and Trofimenkoff, F., "Modeling Operational Amplifiers for Computer-Aided Circuit Analysis," *IEEE Transactions on Circuits and Systems*, vol. 18, no. 1, 1971, pp. 205–207.
- Widlar, R. J., "A New Breed of Linear ICs runs at 1-volt levels," *Electronics*, March 29, 1979, pp. 115–119.
- , "DC Error Reduction in Bipolar Opamps," *1980 IEEE Solid State Circuits Conference*, vol. 23, February 1980, pp. 204–205.
- , "Design Techniques for Monolithic Operational Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 4, no. 4, August 1969, pp. 184–191.
- , "Low Voltage Techniques [for Micropower Operational Amplifiers]," *IEEE Journal of Solid-State Circuits*, vol. 13, no. 6, December 1978, pp. 838–846.
- , "Some Circuit Design Techniques for Linear Integrated Circuits," *IEEE Transactions on Circuit Theory*, vol. CT-12, no. 4, December 1965, pp. 586–590.
- Widlar, R. J., Dobkin, R., and Yamatake, M., "New Op Amp Ideas," *National Semiconductor Application Note 211*, December 1978.
- Widlar, R. J., and Yamatake M., "A 150W Opamp," *1985 IEEE Digest of Technical Papers, Solid State Circuits Conference*, vol. 27, February 1985, pp. 140–141.
- Wooley, B. A., and Pederson, D. O., "A Computer-Aided Evaluation of the 741 Amplifier," *IEEE Journal of Solid-State Circuits*, vol. 6, no. 6, December 1971, pp. 357–366.
- Yang, H. C., and Allsot, D. J., "Considerations for Fast Settling Operational Amplifiers," *IEEE Transactions on Circuits and Systems*, vol. 37, no. 3, March 1990, pp. 326–334.

Review of Current Feedback Operational Amplifiers

In This Chapter

- ▶ Current-feedback operational amplifiers are devices where the main feedback path from output to input carries a current, and not a voltage, as in voltage feedback op-amps. Current-feedback op-amps do not suffer from the gain bandwidth product paradigm as do voltage feedback operational amplifiers, and can achieve very high bandwidths. However, they do have other limitations.

Conventional Voltage-Feedback Op-Amp and the Constant “Gain Bandwidth Product” Paradigm

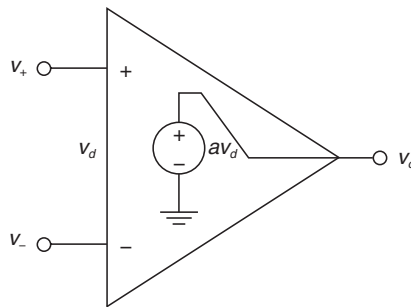
A conventional voltage-feedback operational amplifier is shown in **Figure 13-1**. Let’s assume that the op-amp is dominated by a single pole at a very low frequency. We can express the op-amp open-loop transfer function under this assumption as:¹

$$a(s) \approx \frac{A_o}{\tau s + 1} \approx \frac{K}{s} \quad [13-1]$$

where A_o is the DC gain of the amplifier and $1/\tau$ is the pole frequency of the op-amp dominant pole. Note that we’re approximating the op-amp as behaving as an integrator for frequencies of interest. This simplifies the mathematics in the following discussion considerably.

¹ Since we’re interested in the detail of the high-frequency breakpoint, we’ll make the simplifying assumption that the low-frequency pole is at or near DC. This simplifies the math significantly. In a generic op-amp there’s a dominant pole at a very low frequency and higher-frequency poles that can affect closed-loop stability. The dominant pole is at approximately g_m/C_c , where g_m is the transconductance of the input stage and C_c is the compensating capacitor value. For typical op-amps, this dominant pole is at a few hertz.

Figure 13-1: Conventional voltage-feedback op-amp.
The output voltage is
 $v_o = av_d = a(v_+ - v_-)$.

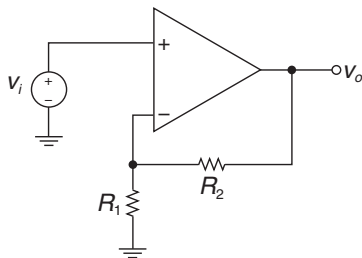


Let's put this amplifier into a positive gain configuration of **Figure 13-2a**. The ideal low-frequency closed-loop gain of this amplifier is:

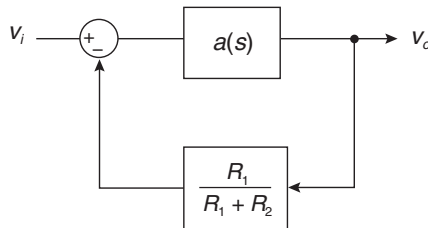
$$G = 1 + \frac{R_2}{R_1} = \frac{R_1 + R_2}{R_1} \quad [13-2]$$

Using the block diagram of **Figure 13-2b**, let's figure out the overall transfer function of this amplifier:

$$\frac{v_o}{v_i} = \frac{\frac{K}{s}}{1 + \left(\frac{K}{s}\right)\left(\frac{R_1}{R_1 + R_2}\right)} = \frac{K}{s + \frac{K}{G}} = \frac{G}{\frac{G}{K}s + 1} \quad [13-3]$$



(a)

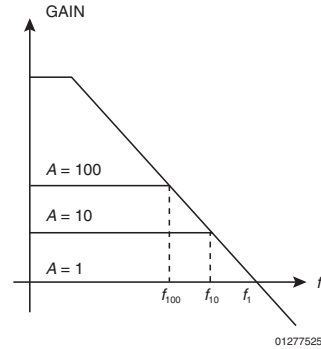


(b)

Figure 13-2: Conventional voltage-feedback op-amp.
(a) Configured as positive gain amplifier. (b) Block diagram.

Note that this amplifier has a low-frequency gain of G (as expected) and a bandwidth of (K/G) . So, if the closed-loop gain goes up, the closed-loop bandwidth goes down. This is the gain-bandwidth product paradigm for a voltage-feedback amplifier, which is illustrated in **Figure 13-3**. As we'll see, the current feedback op-amp is not subject to this same gain-bandwidth product limitation.

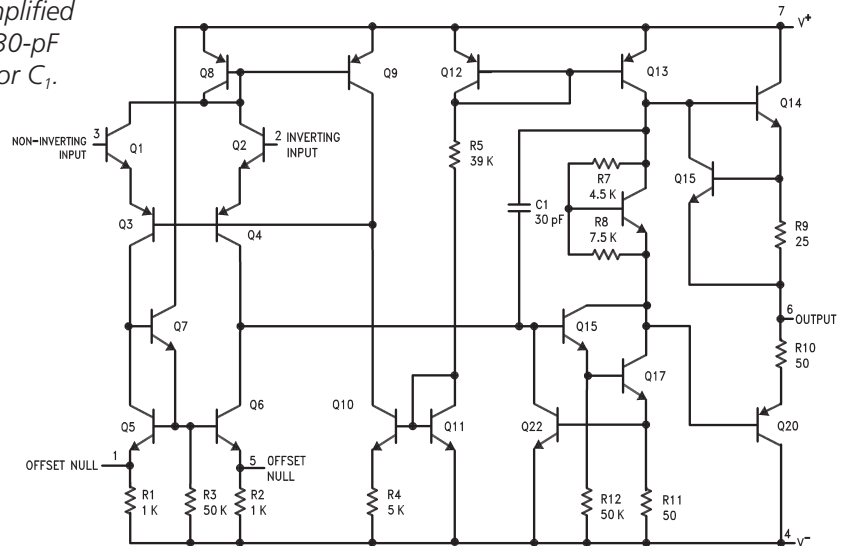
Figure 13-3: Illustration of gain-bandwidth product for conventional voltage-feedback amplifier. As the closed-loop gain is increased, the closed-loop bandwidth decreases.



Slew Rate Limitations in Conventional Op-Amps

In a conventional operational amplifier, a compensating capacitor is used to tailor the open-loop frequency response so that it is dominated by a single pole. In a conventional internally compensated op-amp (such as the LM741 in **Figure 13-3** below) there is a compensating capacitor internal to the device. The maximum rate at which we can charge and discharge this capacitor (and hence the output voltage) depends on the bias current for the previous differential stage (provided originally by Q_8). The specification for slew rate for this op-amp is $\sim 0.5 \mu\text{s}$, consistent with an input bias current of $\sim 15 \mu\text{A}$.²

Figure 13-4: LM741 simplified schematic.³ Note the 30-pF compensation capacitor C_1 .



² The slew rate calculation is simple enough: $dV/dt = I/C$ where I is the input stage bias current and C is the compensation capacitor value. A detailed look at the inner workings of the LM741 can be found in Gray, Hurst Lewis and Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th edition, Chapter 6.8. It is shown in this reference that $I \sim 19 \mu\text{A}$ which results in a slew rate of $\sim 0.6 \text{ V}/\mu\text{sec}$ for a 30-pF compensating capacitor.

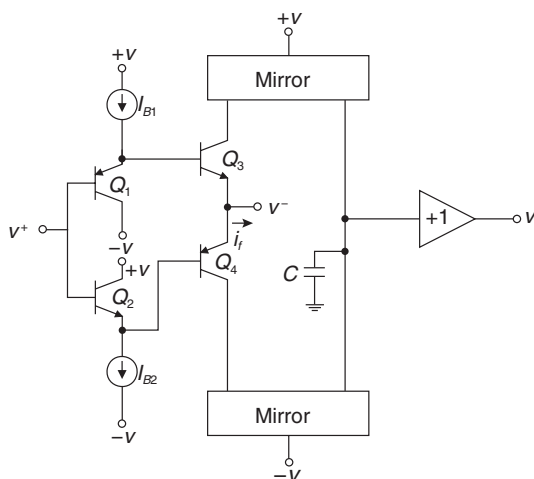
³ This particular datasheet is from National Semiconductor; other vendors make this op-amp as well. Reprinted with permission of National Semiconductor.

Basic Current Feedback Op-Amp

The basic current-feedback op-amp (**Figure 13-5**) differs from the conventional voltage-feedback op-amp in several respects. The negative input (v_-) has a very low input impedance (by design), since this negative node is the place where we'll be feeding back current. A unity-gain buffer (Q_1 – Q_3 and Q_2 – Q_4) forces the v_- input to follow the v_+ input.⁴

Since we use negative *current* feedback to the inverting v_- input,⁵ we want low output impedance at this v_- node, so that the input impedance won't significantly affect the amount of current that is fed back. The input amplifier Q_3/Q_4 and associated current mirrors convert the feedback current i_f flowing out of the negative input to a voltage at the high-gain node at the output of the current mirrors. This voltage at the high-gain node is further buffered by a unity-gain buffer to the output.

Figure 13-5: One topology of the current-feedback opamp.



Note that in **Figure 13-5** no voltage amplification occurs inside the amplifier. The input followers Q_1 – Q_2 – Q_3 – Q_4 have a voltage gain of +1. The current mirrors provide a current gain of +1 with very little voltage gain. At the high-gain node the mirror current is converted to a voltage, which is buffered to the output. Therefore, there is no “Miller effect” in the current-feedback amplifier.

⁴ This type of 4-transistor follower is sometimes called a “diamond buffer” or a “diamond follower.” See, e.g., Burr-Brown (later, Texas Instruments) application note AB-181 by K. Lehmann, “Diamond Transistor OPA660,” and W. Lillis and A. Wang, “Complementary current mirror for correcting input offset voltage of diamond follower, especially as input stage for wideband amplifier,” U.S. Patent # 4,893,091, issued January 9, 1990.

⁵ The concept of current feedback is not new; designs for tube amplifiers using cathode feedback date back to the 1920s.

In order to analyze approximately the gain and bandwidth of the current-feedback op-amp, let's consider a simplified model that captures the first-order effects. **Figure 13-6** shows this current-feedback op-amp reduced to a simplified form. The input resistance R_{in} is the parallel combination of the output resistances of Q_3 and Q_4 .⁶

$$R_{in} \approx \frac{1}{g_{m3}} \parallel \frac{1}{g_{m4}} \quad [13-4]$$

For a Q_3/Q_4 bias current of 250 microamps, this input resistance is approximately 52Ω .⁷

To first-order, this amplifier has a single pole due to the $R_T C_T$ time constant of the high-gain node. The resistance R_T at the high-gain node is a combination of the output resistances of the high-side and low-side current mirrors. Assuming that we have cascoded or Wilson current mirrors (to increase the output resistance) the resistance at the high-gain node is:

$$R_T \approx \frac{r_{\mu,npn}}{2} \parallel \frac{r_{\mu,npn}}{2} \quad [13-5]$$

Assuming a 250- μ A mirror bias current and base-width modulation factors $\eta_{npn} = \eta_{pnp} = 10^{-3}$, and with $\beta_{Fnpn} = \beta_{Fpnp} = 100$, this high-gain node resistance $R_T = 2.6 \text{ M}\Omega$.⁸

The capacitor C_T is set by device parameter and layout parasitics.

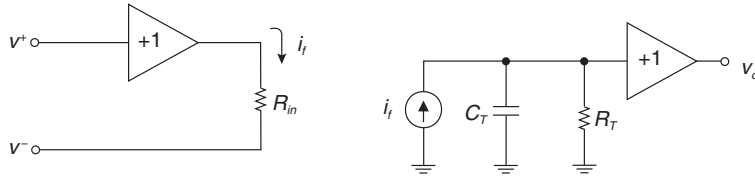


Figure 13-6: Equivalent diagram of current-feedback op-amp.
The feedback current i_f is mirrored to the high gain node.

If we're interested in DC accuracy, the DC voltage gain of this amplifier is of importance. The open loop voltage gain is:

$$a_{v,ol} \approx \frac{R_T}{R_{in}} \approx \frac{2.6 \times 10^6}{52} \approx 50,000 \quad [13-6]$$

⁶ This simplified calculation ignores the base-spreading resistance r_x .

⁷ Remember that the input resistance looking into the emitter of a transistor in the forward active region is approximately $1/g_m$, if the base resistance r_x and the source resistance seen from the base terminal are neglected. The input resistance, in close form, is:

$$r_{in,emitter} \approx \frac{R_s + r_x + r_\pi}{h_{fe} + 1} \approx \frac{1}{g_m} \text{ if } R_s, r_x \ll r_\pi$$

⁸ These numbers are consistent with the Analog Devices AD844, which lists a nominal input resistance at the inverting input of 50Ω , and a high-gain node resistance of $3 \text{ M}\Omega$.

Chapter 13

A current-feedback amplifier connected in a positive-gain configuration is shown in **Figure 13-7**. A block diagram of this system is shown in **Figure 13-8**.

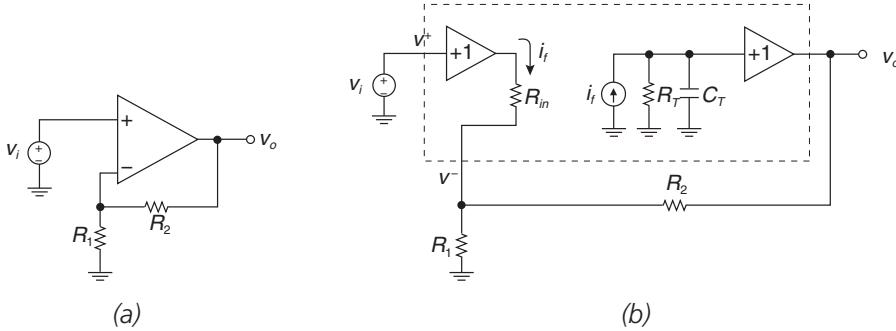


Figure 13-7: Current-feedback op-amp configured for positive gain. (a) Circuit. (b) Model.

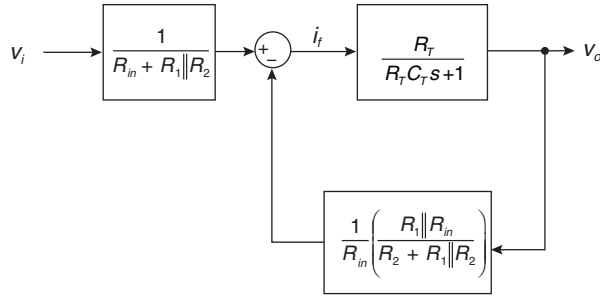


Figure 13-8: Current-feedback op-amp configured for positive closed-loop gain, block diagram.

The overall gain for this amplifier is:

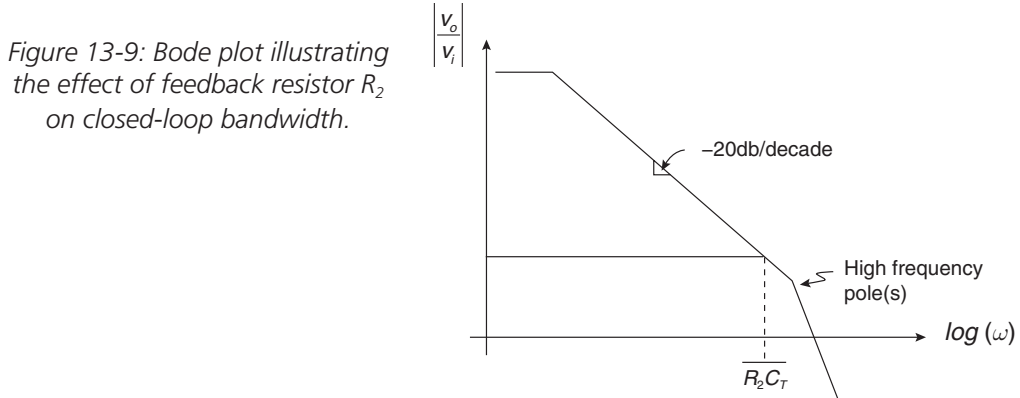
$$\frac{v_o}{v_i} = \left(\frac{1}{R_{in} + R_1 \parallel R_2} \right) \left(\frac{\frac{R_T}{R_T C_T s + 1}}{1 + \left(\frac{R_T}{R_T C_T s + 1} \right) \left(\frac{1}{R_{in}} \right) \left(\frac{R_1 \parallel R_{in}}{R_2 + R_1 \parallel R_{in}} \right)} \right) \quad [13-7]$$

Although this looks like a mess, we can simplify this by recognizing that the input resistance to the unity-gain input buffer is very small ($R_{in} \ll R_1$ and R_2) and that the resistance at the high-gain node is very large (or $R_T \gg R_1$ and R_2). Using these assumptions, we arrive at the gain of this amplifier as:

$$\frac{v_o}{v_i} \approx \left(\frac{R_1 + R_2}{R_1 R_2} \right) \left(\frac{R_T}{R_T C_T s + 1 + \frac{R_T}{R_2}} \right) \approx \left(\frac{R_1 + R_2}{R_1} \right) \left(\frac{1}{R_2 C_T s + 1 + \frac{R_2}{R_T}} \right) \approx G \left(\frac{1}{R_2 C_T s + 1} \right) \quad [13-8]$$

Note that the ideal closed-loop gain G is the same as with the voltage-feedback amplifier. Also, note the absence of gain-bandwidth product in this case. The overall bandwidth does depend on the value of the feedback capacitor R_2 but we can independently set the closed-loop gain (by varying R_1) without affecting bandwidth.⁹

Next, we'll comment on stability. We note that the bandwidth as shown before is roughly $1/(R_2 C_T)$, ignoring high-frequency poles, as shown in **Figure 13-9**. We also note that there are indeed higher frequency poles unaccounted for in this simplified model. Therefore, manufacturers will specify a minimum feedback resistance¹⁰ requirement to ensure stability.



Absence of Slew Rate Limit in Current Feedback Op-Amps

In the current feedback op-amp we do not have the same slew rate limit as in a conventional voltage-feedback op-amp. Note that in the current-feedback op-amp, the current injected into the negative (v_-) input to the input stage gets mirrored and eventually charges and discharges the capacitance C_T at the high-gain node. Therefore, for a large voltage swing at the output of the amplifier, there is more current fed back to the v_- terminal. Following this line of reasoning, there is no intrinsic slew rate limit in a current-feedback op-amp.¹¹

⁹ Note that this works to first order only. Second-order effects do come into play to limit the bandwidth as well.

¹⁰ Note that if you try to push up the bandwidth by reducing R_2 , eventually you will “run into” the high-frequency poles, creating a potential stability problem.

¹¹ There are, of course, second-order limitations that affect the ultimate maximum slew rate of current feedback op-amps. There is some parasitic capacitance at the high-gain node. However, a survey of some current feedback op-amps shows that the slew rate is indeed high. See, e.g., Analog Devices’s AD844 (2000 V/ μ s); National Semiconductor’s LM6181 (2000 V/ μ s); Texas Instruments’s TMS3110 (1300 V/ μ s); Linear Technology LT1227 (1100 V/ μ sec). There are lots more.

Example 13.1: An admittedly very crude current-feedback op-amp discrete design

Let's look at an illustrative design of a discrete current-feedback op-amp circuit and simulate with 2N3904 and 2N3906 transistors. Note that this is by no means an optimized design; the resultant design is likely to have problems with voltage offset, transient response and the like. Rather, this simple example illustrates many of the concepts that we've gone over in the previous analytic discussion.

Shown in **Figure 13-10** is the design, configured initially for unity feedback. At first, assume that the value for R_1 is 1 M Ω . In this configuration, we expect approximately unity gain, since, as we've shown earlier, the closed-loop gain (at low frequencies) is approximately:

$$G = 1 + \frac{R_2}{R_1} = \frac{R_1 + R_2}{R_1} \quad [13-9]$$

Going through the design, we see the following functional blocks:

- Q_1 – Q_4 is the input buffer, providing high input impedance at v_+ (the bases of Q_1 and Q_3) and a low input impedance at v_- (the emitters of Q_2 and Q_4). Q_1 and Q_3 are biased by current sources, and the collector currents of Q_2 and Q_4 are approximately:

$$I_{C2} \approx I_{C4} \approx \sqrt{I_{C1}I_{C3}} \quad [13-10]$$

We'll see in a later section in this book (on the translinear principle) how this calculation is done.

- Q_5/Q_7 and Q_6/Q_8 are current mirrors that mirror the collector currents of Q_2 and Q_4 to the high gain node. These mirrors act as a current buffer, mirroring the feedback current to the high gain node. The high gain node is the collector connection of Q_7 and Q_8 . The output resistance at this node is the parallel combination of the output resistances of current mirror transistors Q_7 and Q_8 ; the total resistance at this node to ground is this current mirror resistance in parallel with the input resistance of the following stage Q_9/Q_{10} .
- Q_9/Q_{10} is a unity-gain push-pull output stage, which provides a low impedance at the output v_{out} node.
- R_2 is the feedback resistor, and R_1 is the input resistor at the inverting input v_- . Note that our previous analysis of current-feedback op-amps showed that the main bandwidth limitation is our selection of feedback resistor R_2 .

The SPICE bias point values show some interesting results. For instance, the input buffer (Q_1 – Q_4) does have a voltage offset as configured; note that while the input is at 0V, the negative input (v_-) is at –12.9 mV. This voltage offset cascades down through the rest of the amplifier, resulting in an output voltage of –681 mV when the v_+ input is at zero potential.

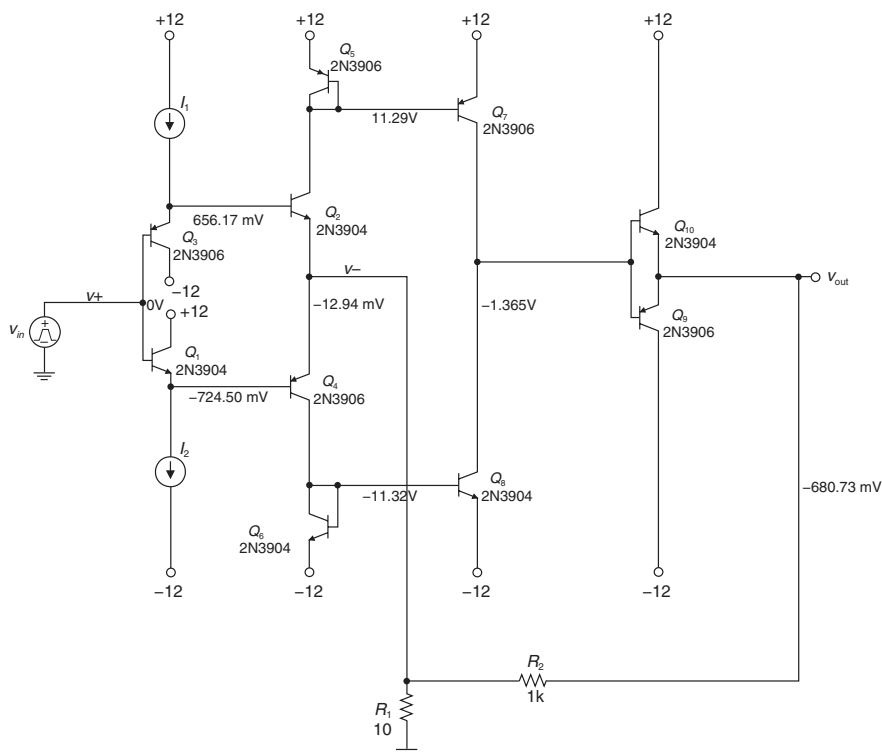


Figure 13-10: Discrete-design current-feedback op-amp configured for positive gain. Node voltages shown are predicted by PSPICE.

Results for AC analysis of this op-amp for values $R1 = 1\text{M}\Omega$, 100Ω and 10Ω are shown in **Figure 13-11**. Note that the closed-loop bandwidth is roughly constant over a wide range of closed-loop gains.

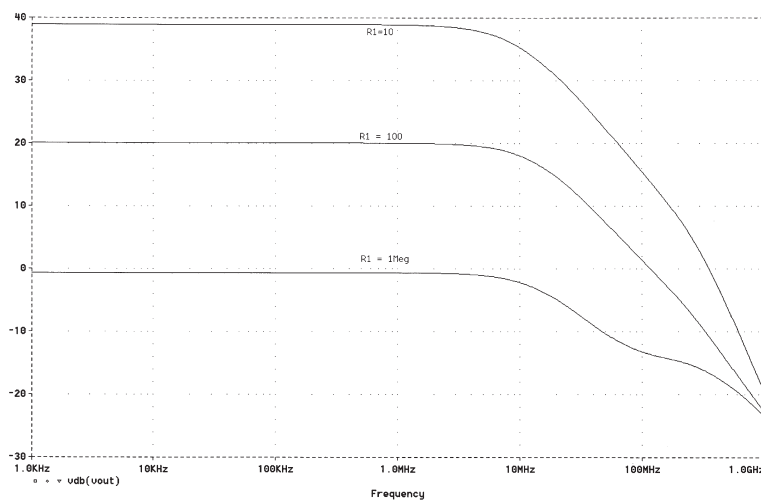


Figure 13-11: PSPICE results for AC analysis of discrete CF op-amp.

In **Table 13-1** is a summary of PSPICE results comparing the ideal gain with SPICE simulated, and resultant bandwidth and amplifier gain-bandwidth product. Note that there are some minor differences in calculated gains vs. PSPICE gains, due in part to loading effects not accounted for in the simplified analyses.

Table 13-1: PSPICE results for current-feedback op-amp design example.

R_1	Ideal gain	Ideal gain (dB)	PSPICE gain	PSPICE gain (dB)	PSPICE small-signal bandwidth	GBP
1 M Ω	1.001	0.009	0.93	-0.6	15 MHz	14 MHz
100 Ω	11.0	20.8	10.1	20.1	12.7 MHz	128.5 MHz
10 Ω	101	40.1	88.1	39	8.7	766.5 MHz

The step response for the gain-of-101 configuration is shown in **Figure 13-12**. The input step was set to 1 mV. The risetime is approximately 41 ns, consistent with our small-signal bandwidth of 8.7 MHz.¹² Note the poor voltage offset of this amplifier.¹³

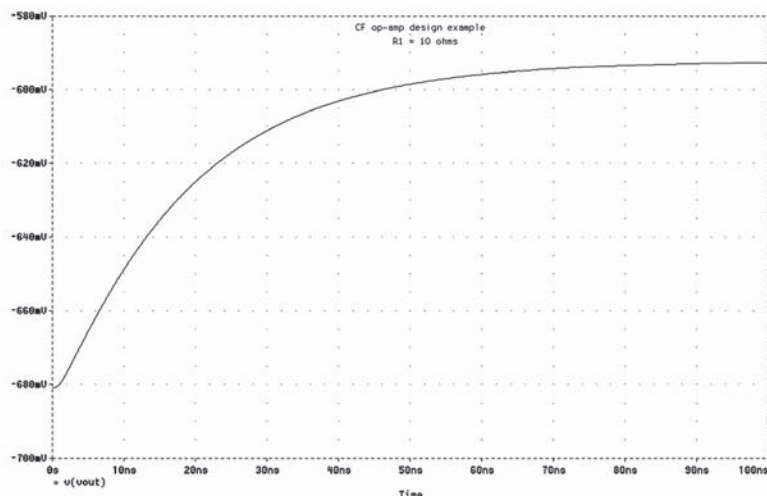


Figure 13-12: PSPICE results for step response of discrete current feedback op-amp.


¹² The input step needs to be small enough so that the operating point of the transistors doesn't significantly change and the amplifier doesn't slew rate limit, hence the use of the small 1-mV input step. Remember our calculation for the risetime of a first-order system, where it was found that $\text{risetime} = 0.35/\text{bandwidth}$, where bandwidth is in hertz. This op-amp has more than one pole (more than a dozen in this example). However, the amplifier small-signal AC response and small-signal step response is indeed dominated by a single pole, due to loading at the high-gain node.

¹³ This is why the output starts at -680 mV (instead of zero). In a real-world CF op-amp, extra biasing circuitry is added to reduce this voltage offset significantly.

Manufacturer's Datasheet Information for a Current Feedback Amplifier

Let's have a look at excerpts of a manufacturer's datasheet. A section of the front page of the datasheet of the National LM6181 is shown in **Figure 13-13**. Note the high slew rate (2000 V/ μ s). This op-amp may typically be used in video amplifier applications where high bandwidth, high slew rate, and low differential phase and gain are needed.

The simplified block diagram is shown in **Figure 13-14**. There is a four-transistor input buffer providing the positive and negative op-amp inputs. The feedback current is mirrored through a cascade current mirror. Cascoding provides higher current source output impedance at the high-gain node, and hence more open-loop voltage gain. The voltage at the high-gain node is buffered to the output by another four-transistor unity-gain buffer. Note that the output includes short-circuit protection; if the voltage across the output resistors R exceeds approximately 0.6V, the output transistors are shut down.


May 1998

LM6181

100 mA, 100 MHz Current Feedback Amplifier

General Description

The LM6181 current-feedback amplifier offers an unparalleled combination of bandwidth, slew-rate, and output current. The amplifier can directly drive up to 100 pF capacitive loads without oscillating and a 10V signal into a 50 Ω or 75 Ω back-terminated coax cable system over the full industrial temperature range. This represents a radical enhancement in output drive capability for an 8-pin DIP high-speed amplifier making it ideal for video applications.

Built on National's advanced high-speed VIP™ II (Vertically Integrated PNP) process, the LM6181 employs current-feedback providing bandwidth that does not vary dramatically with gain; 100 MHz at $A_v = -1$, 60 MHz at $A_v = -10$. With a slew rate of 2000V/ μ s, 2nd harmonic distortion of -50 dBc at 10 MHz and settling time of 50 ns (0.1%) the LM6181 dynamic performance makes it ideal for data acquisition, high speed ATE, and precision pulse amplifier applications.

Features

(Typical unless otherwise noted)

- Slew rate: 2000 V/ μ s
- Settling time (0.1%): 50 ns
- Characterized for supply ranges: ± 5 V and ± 15 V
- Low differential gain and phase error: 0.05%, 0.04°
- High output drive: ± 10 V into 100 Ω
- Guaranteed bandwidth and slew rate
- Improved performance over EL2020, OP160, AD844, LT1223 and HA5004

Applications

- Coax cable driver
- Video amplifier
- Flash ADC buffer
- High frequency filter
- Scanner and Imaging systems

Figure 13-13: National Semiconductor LM6181 current feedback op-amp.
Reprinted with permission of National Semiconductor Corporation.

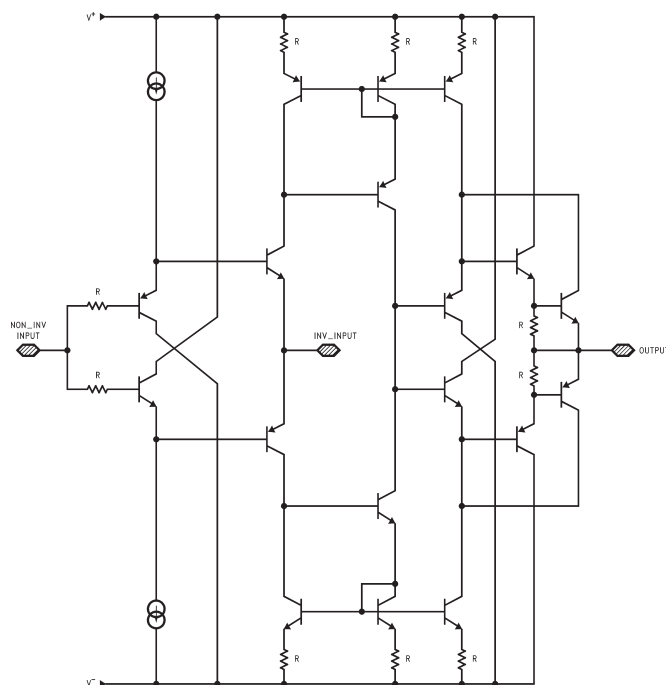


Figure 13-14: National Semiconductor current feedback op-amp simplified diagram. Reprinted with permission of National Semiconductor Corporation.

The closed-loop frequency response curves (**Figure 13-15**) show that the bandwidth is roughly constant over a wide gain setting from -1 to -10 .

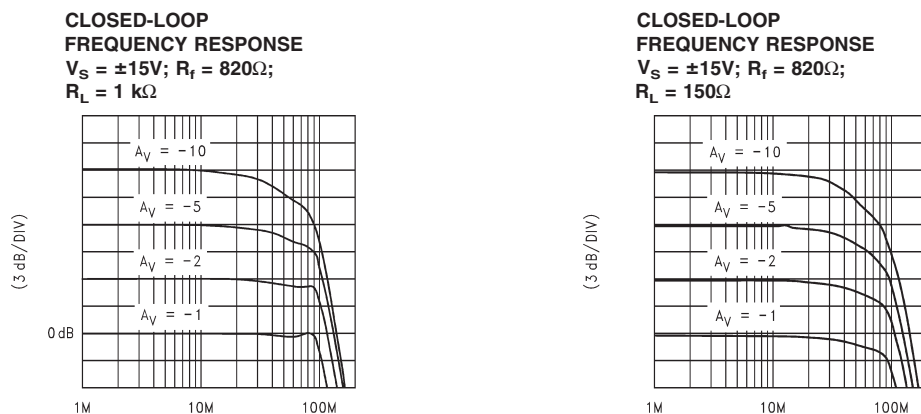


Figure 13-15: National Semiconductor LM6181 current feedback op-amp small-signal bandwidth. Reprinted with permission of National Semiconductor Corporation.

A More Detailed Model and Some Comments on Current-Feedback Op-Amp Limitations

A more detailed model of the current-feedback op-amp, modeling some further effects that may affect bandwidth, gain and stability is shown in **Figure 13-16**. For instance, C_{in-} models the parasitic capacitance at the inverting input node. This capacitance is due to internal transistor capacitances as well as any external parasitic capacitances due to bond wires and IC leads. The resistance r_o models the small (but finite) output resistance of the output unity-gain buffer. After using the simplified equivalent circuit model (**Figure 13-6**) for initial analysis, it often behooves the designer to take into account these additional circuit components that will affect circuit operation as well.

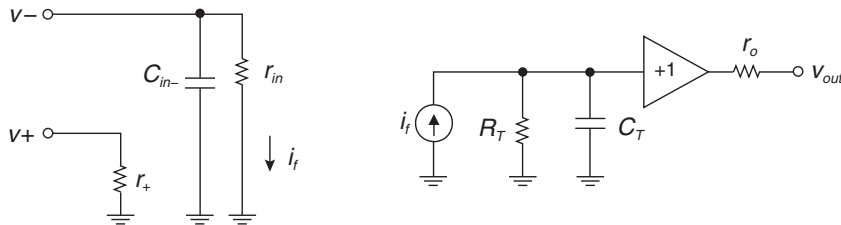


Figure 13-16: A more detailed model of the current-feedback op-amp. Parasitics that have been added to the previous model include: r_+ (input resistance at the positive input terminal); C_{in-} (input capacitance at the negative input terminal) and r_o (output resistance).

There are other design issues to take into account when using a current-feedback op-amp, as compared to a voltage feedback op-amp.

DC performance. Operation of the current feedback op-amp with acceptable DC performance (i.e., voltage offset, CMRR, etc.) relies on matching between NPN and PNP transistors in the front end. Further degradation in the voltage offset is caused by the Early voltage of the transistors in the input buffer.

Low impedance at inverting input. Many current feedback amplifiers, by their nature, rely on a low impedance at the inverting input.

Minimum value of feedback resistor. Due to stability issues, a minimum value of feedback resistor must be used in closed-loop configurations to ensure stability.

Chapter 13 Problems

Problem 13.1

Using a conventional op-amp model, draw a circuit using the 741 operational-amplifier providing a gain of +100. Find the bandwidth of this amplifier circuit.

Problem 13.2

Using the AD844 datasheet, extract current-feedback op-amp parameters and fill in the equivalent circuit model of **Figure 13-6**.

Problem 13.3

Using open-circuit time constants, estimate the bandwidth of the four-transistor input circuit (Q_1 – Q_4) of **Figure 13-5**. Assume that Q_3 and Q_4 are each biased at a collector current I_o and that each transistor has a current gain-bandwidth product of $\omega_T = 6.3 \times 10^9$ radians/second. Assume that each transistor has $C_\mu = 0.5$ pF, and assume that $I_{B1} = I_{B2} = I_o = 1$ mA. Make other assumptions as needed and state them.

Problem 13.4

Using the circuit of **Figure 13-5**, find the collector current in transistors Q_3 and Q_4 , assuming that the input transistors Q_1 and Q_2 are biased as shown. Assume that $i_f = 0$ and that the input v_+ is grounded for purposes of this calculation. Furthermore, assume that the reverse saturation current is $I_{S,NPN}$ for Q_2 and Q_3 , and is $I_{S,PNP}$ for Q_1 and Q_4 .

References

- Bowers, D. F., "The so-called current-feedback operational amplifier-technological breakthrough or engineering curiosity?" *IEEE International Symposium on Circuits and Systems, ISCAS '93*, May 3–6 1993, pp. 1054–1057.
- Bowers, D. F., and Wurcer, S. A., "Recent developments in bipolar operational amplifiers," *Proceedings of the 1999 Bipolar/BiCMOS Circuits and Technology Meeting*, September 26–28, 1999, pp. 38–45.
- Franco, S., "Current Feedback Amplifiers Benefit High-Speed Designs," *EDN*, January 5, 1989, pp. 161–172; "Analytical foundations of current-feedback amplifiers," *1993 IEEE International Symposium on Circuits and Systems, (ISCAS '93, vol.2)* 3–6 May 3–6 1993, pp. 1050–1053.
- Harvey, B., "Current feedback opamp limitations: a state-of-the-art review," *1993 IEEE International Symposium on Circuits and Systems, (ISCAS '93)*, May 3–6 1993, p. 1066–1069.
- , "Dual-Amplifier Designs Increase the Accuracy of Current-Feedback Amps," *EDN*, December 9, 1993, pp. 129–134.
- , "Take Advantage of Current-Feedback Amps for High-Frequency Gain," *EDN*, March 18, 1993, pp. 215–222.
- Intersil, Inc., "Current Feedback Amplifier Theory and Applications," *Application Note # AN9420.1*, April 1995.
- , "An Intuitive Approach to Understanding Current Feedback Amplifiers," *Application Note # AN9787*, February 1998.
- , "Converting From Voltage-Feedback to Current-Feedback Amplifiers," *Application Note # AN9663.1*, April 1999.
- , "A Designer's Guide for the HA-5033 Video Buffer," *Application Note AN548.1*, November 1996.
- Koullias, I. A., "A Wideband Low-Offset Current-Feedback Op Amp Design," *Proceedings of the 1989 Bipolar Circuits and Technology Meeting*, Minneapolis, September 18–19, 1989, pp. 120–123.
- Lidgey, F. J., and Hayatleh, K., "Current-feedback operational amplifiers and applications," *Electronics & Communication Engineering Journal*, vol. 9, no. 4, Aug. 1997, pp. 176–182.
- Lillis, William, and Wang, Anthony, "Complementary Current Mirror for Correcting Input Offset Voltage of Diamond Follower, Especially as Input Stage for Wide-Band Amplifier," U.S. Patent # 4,893, 091, filed October 11, 1988; issued January 9, 1990.
- Mancini, Ronald, "Converting From Voltage-Feedback to Current-Feedback Amplifiers," *Electronic Design, (Analog Applications Issue)*, June 26, 1995, pp. 37–46.
- National Semiconductor, "Current Feedback Op Amp Applications Circuit Guide," *Application Note # OA-07*, May 1988.
- , "Frequent Faux Pas in Applying Wideband Current Feedback Amplifiers," *Application Note # OA-15*, August 1990.

- , “Current-Feedback Myths Debunked,” Application Note # OA-20, July 1992.
- , “Current Feedback Amplifiers,” Application Note # OA-31, November 1992.
- , “Current Feedback Loop Gain Analysis and Performance Enhancement,” Application Note # OA-13, January 1993.
- , “Stability Analysis of Current Feedback Amplifiers,” Application Note # OA-25, May 1995.
- , “Current vs. Voltage Feedback Amplifiers,” Application Note # OA-30, January 1998.
- Palumbo, G., “Current feedback amplifier: stability and compensation,” *Proceedings of the 40th Midwest Symposium on Circuits and Systems*, August 3-6, 1997, pp. 249-252.
- Palumbo, G., and Pennisi, S., “Current-feedback amplifiers versus voltage operational amplifiers,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 48, no. 5, May 2001, pp. 617-623.
- Sauerwald, M., “Current feedback and voltage feedback. The choice amplifiers: which one to use, when and why,” *Northcon/94 Conference Record*, October 11-13, 1994, pp. 285-289.
- Smith, Doug, “Current-feedback amps enhance active-filter speed and performance,” *EDN*, September 17, 1990, pp. 167-172.
- Smith, Douglas, “High Speed Operational Amplifier Architectures,” *Proceedings of the 1993 IEEE Bipolar Circuits and Technology Meeting*, pp. 141-148.
- Smith, Doug, Koen, Mike, and Witulski, Arthur, “Evolution of High-Speed Operational Amplifier Architectures,” *IEEE Journal of Solid-State Circuits*, vol. 29, no. 10, October 1994, pp. 1166-1179.
- Tammam, A. A., Hayatleh, K., Hart, B., and Lidgley, F. J., “High Performance Current-Feedback Op-Amps,” *Proceedings of the 2004 International Symposium on Circuits and Systems, (ISCAS '04)*, May 23-26, 2004, pp. 825-828.
- Texas Instruments, Inc., “Voltage Feedback vs. Current Feedback Op Amps Applications Report,” *TI Application Note*, November 1998.
- Toumazou, Chris, Payne, Alison, and Lidgley, John, “Current-Feedback Versus Voltage Feedback Amplifiers: History, Insight and Relationships,” *1993 IEEE International Symposium on Circuits and Systems, (ISCAS '93)*, May 3-6, 1993, pp. 1046-1049.
- Wilson, B., “Current-mode amplifiers,” *IEEE International Symposium on Circuits and Systems*, May 8-11, 1989, pp. 1576-1579.
- Wong, James, “Current-Feedback Op Amps Extend High Frequency Performance,” *EDN*, October 26, 1989, pp. 211-216.

Analog Low-Pass Filters

In This Chapter

- The basics of analog low-pass filtering are discussed. These techniques are useful for designing analog filters as well as prototypes used for digital filters. This chapter is by no means an all-inclusive tutorial on analog circuit design; rather it is introductory in nature and the reader is referred to other texts for more details, if necessary.

Introduction

The low-pass filter is a ubiquitous component in many different kinds of signal-processing systems. Channel-separation, A/D antialiasing, and general signal processing are applications for low-pass filters, just to name a few. Even if you are a digital filter designer, it behooves you to know something about analog filter design, since many digital filters begin as analog prototypes, and are then transformed to the digital domain. Also, DSP systems generally have an analog front end that includes an analog low-pass filter for antialiasing purposes (**Figure 14-1**).

As with any kind of design, the “devil is in the details.” Your specification will lead you to choices in filter topology and filter order, depending on the attenuation you need, the ripple that you can live with, and also the group delay variation that you can live with. In the following sections, we discuss design issues associated with low-pass filter design. The results can be extended without much trouble to band-pass and high-pass filters as well.

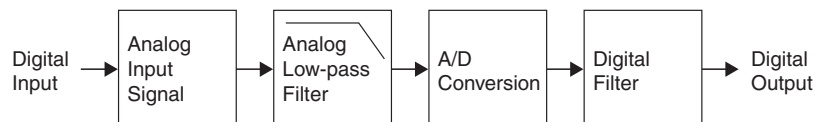


Figure 14-1: Typical digital signal-processing chain.

Review of Low-Pass Filter Basics

The magnitude response of the ideal low-pass filter is shown in **Figure 14-2a**. The gain of this filter is perfectly flat in the passband (for frequencies less than the filter cutoff¹ frequency ω_h), and the response drops to zero for frequencies higher than the cutoff frequency.

The magnitude response of a real-world low-pass filter is shown in **Figure 14-2b**. The non-ideal effects include:

- Possible ripple in the filter passband.
- Possible maximum attenuation floor in the stopband.
- Finite transition width between passband and stopband.

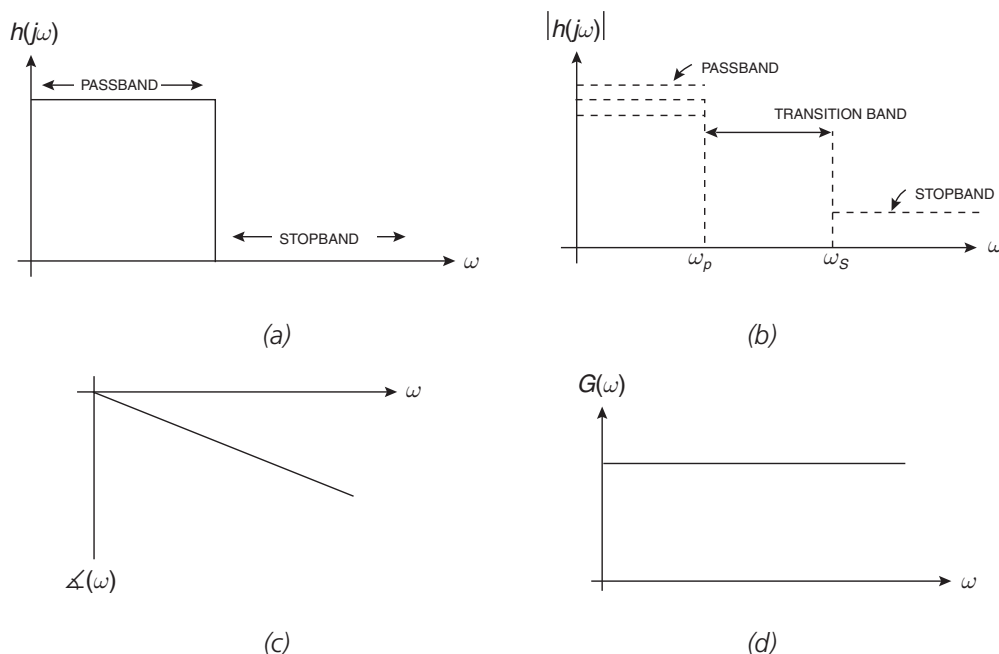


Figure 14-2: Response of low-pass filter. (a) Ideal low-pass filter magnitude $|H(j\omega)|$. (b) Real-world magnitude response $|H(j\omega)|$ showing possible ripple in the passband and stopband, and a finite transition width between passband and stopband. (c) Ideal low-pass filter phase response $\angle(\omega)$ showing a negative phase shift that increases linearly with frequency. (d) Ideal low-pass filter group delay response $G(\omega)$ which is constant.

¹ In general, the cutoff frequency of a filter is specified to be the frequency at which the gain through the filter has dropped to 0.707 of the DC value, or -3dB .

The magnitude response, however, only tells half the story. In addition, we must be concerned with the phase response of filters. As we'll see in the following sections the phase response (and by association the group delay² response) affects the transient response of filters. An ideal filter has a linear phase shift with frequency, and hence constant group delay as in **Figures 14-2c** and **14-2d**. The following section discusses in detail several different low-pass filter types that to varying degrees approximate the ideal magnitude and phase of a low-pass filter.

Butterworth Filter

The Butterworth is a class of filters that provides maximally flat response in the passband. The pole locations for an Nth-order Butterworth filter are found equally spaced around a circle with radius equal to the filter cutoff frequency. A Butterworth filter with –3dB cutoff frequency $\omega_{3dB} = 1$ radian/second has poles at locations:

$$-\sin \frac{(2k-1)\pi}{2N} + j \cos \frac{(2k-1)\pi}{2N} \quad k = 1, 2, \dots, N \quad [14-1]$$

The frequency response is given by:

$$|H(j\omega)| = \frac{1}{\sqrt{1 + \omega^{2N}}} \quad [14-2]$$

The pole transfer functions for Butterworth filters of varying order are shown in **Table 14-1**. Note that the transfer functions have been broken up into first-order and second-order factors.³ Breaking up the transfer function in this fashion will help us to implement our filter, since first- and second-order sections are easily synthesized with op-amps. We also show the transfer function for a filter cutoff frequency of one radian per second.

² The group delay of a filter is given by the derivative with respect to frequency of the phase response, or: $G(\omega) = -\frac{d\angle(\omega)}{d\omega}$. The group delay is a measure of how much a given frequency component is delayed passing through the filter. For low pulse distortion, you want all Fourier components to be delayed by the same amount of time, and hence you want constant group delay, and hence linear phase response.

³ This is done to help facilitate the implementation of the transfer function using op-amps. For instance, first-order sections can be implemented with simple RC filters. Second-order transfer functions can be implemented using any number of op-amp circuits, including the Sallen-Key filter. More on this later on.

Table 14-1: Transfer function for Butterworth filters broken up into first-order and second-order factors. Transfer functions are shown for varying filter order N , with filter cutoff frequency $\omega_{3dB} = 1$ rad/sec.

N	Transfer function
2	$\frac{1}{s^2 + 1.414s + 1}$
3	$\frac{1}{(s + 1)(s^2 + s + 1)}$
4	$\frac{1}{(s^2 + 0.7654s + 1)(s^2 + 1.8478s + 1)}$
5	$\frac{1}{(s + 1)(s^2 + 0.6180s + 1)(s^2 + 1.6180s + 1)}$
6	$\frac{1}{(s^2 + 0.5176s + 1)(s^2 + 1.4142s + 1)(s^2 + 1.9318s + 1)}$
7	$\frac{1}{(s + 1)(s^2 + 0.4450s + 1)(s^2 + 1.2480s + 1)(s^2 + 1.8019s + 1)}$
8	$\frac{1}{(s^2 + 0.3902s + 1)(s^2 + 1.1111s + 1)(s^2 + 1.6629s + 1)(s^2 + 1.9616s + 1)}$

The magnitude response, step response and group delay response for Butterworth filters are shown in **Figure 14-3**, **Figure 14-4** and **Figure 14-5**, respectively. We note the following:

- As the filter order increases, the sharpness of the attenuation characteristic in the transition band increases. For instance, at 10 radians/second, an 8th order filter has an attenuation of roughly 160dB, while lower-order filters have less attenuation.
- With regard to step response, the overshoot and delay through the filter increases as the filter order increases.
- With regard to group delay, the peak-peak variation in group delay increases as the filter order increases.

The preceding allows us to make some general statements regarding the relationship between filter frequency response, group delay, and transient response. A filter with a sharper cutoff will in general have more group delay variation and more overshoot and/or ringing in the frequency response. This is the reason why group delay variation is an important design parameter in filter design: more group delay variation results in more pulse distortion of a waveform passing through a filter. Conversely, filters with near-constant group delay pass pulses without significant distortion.

Figure 14-3: Butterworth filter frequency response for filters with cutoff frequency 1 radian/second, filter orders $N = 2$ to 8.

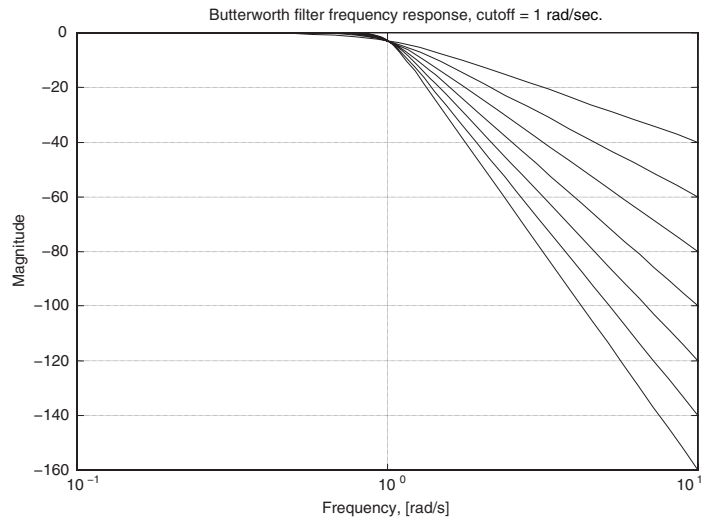


Figure 14-4: Butterworth filter step response for filters with cutoff frequency 1 radian/second, filter orders $N = 2$ to 8.

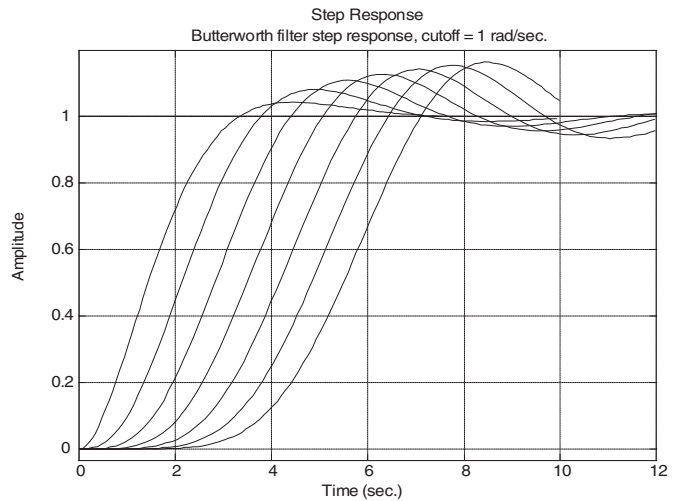
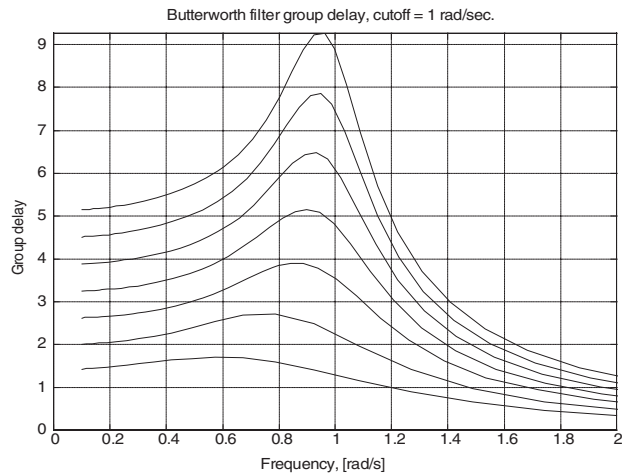


Figure 14-5: Butterworth filter group delay response for filters with cutoff frequency 1 radian/second, filter orders $N = 2$ to 8.



Chebyshev Filter

The Chebyshev filter has some ripple in the passband, and a sharper cutoff in the transition band than the Butterworth filter. As we'll see, the price one pays for using the Chebyshev design compared to the Butterworth is more group delay variation.

The pole locations of an N th order Chebyshev are at locations $\sigma_k + j\omega_k$, with real parts and imaginary parts given by:

$$\begin{aligned}\sigma_k &= -\sinh\left[\frac{1}{N}\sinh^{-1}\left(\frac{1}{\varepsilon}\right)\right]\sin(2k-1)\frac{\pi}{2N}, \quad k=1,2,\dots,N \\ \omega_k &= \cosh\left[\frac{1}{N}\sinh^{-1}\left(\frac{1}{\varepsilon}\right)\right]\cos(2k-1)\frac{\pi}{2N}, \quad k=1,2,\dots,N\end{aligned}\tag{14-3}$$

The Chebyshev filter is characterized by two parameters: filter order N and allowable ripple in the passband. The ripple parameter ε is given by:

$$\varepsilon = \sqrt{10^{\frac{R_{db}}{10}} - 1}\tag{14-4}$$

where R_{db} is the allowable peak ripple in decibels. These pole locations yield a magnitude transfer function that is sharper than that of the Butterworth filter. Shown in **Figure 14-6a** is the Bode plot of an $N = 5$, 0.20dB ripple Chebyshev filter.

Looking at the passband detail of this filter, we plot the frequency response from 0.1 to 1 radian/second in **Figure 14-6b**, we see that the filter order ($N = 5$) equals the number of up and down ripples in the passband. We also note that this Chebyshev formulation does not result in a filter cutoff frequency of exactly 1 radian/second. Rather, this formulation results in a filter where we leave the allowable ripple band at 1 radian/second; the -3 dB point is somewhat higher. We can show that the -3 dB point is given by:

$$\omega_{3db} = \cosh\left[\left(\frac{1}{N}\right)\cosh^{-1}\left(\frac{1}{\varepsilon}\right)\right]\tag{14-5}$$

For instance, for our $N = 5$, 0.20dB example, the -3 dB bandwidth is 1.10 radians/second.

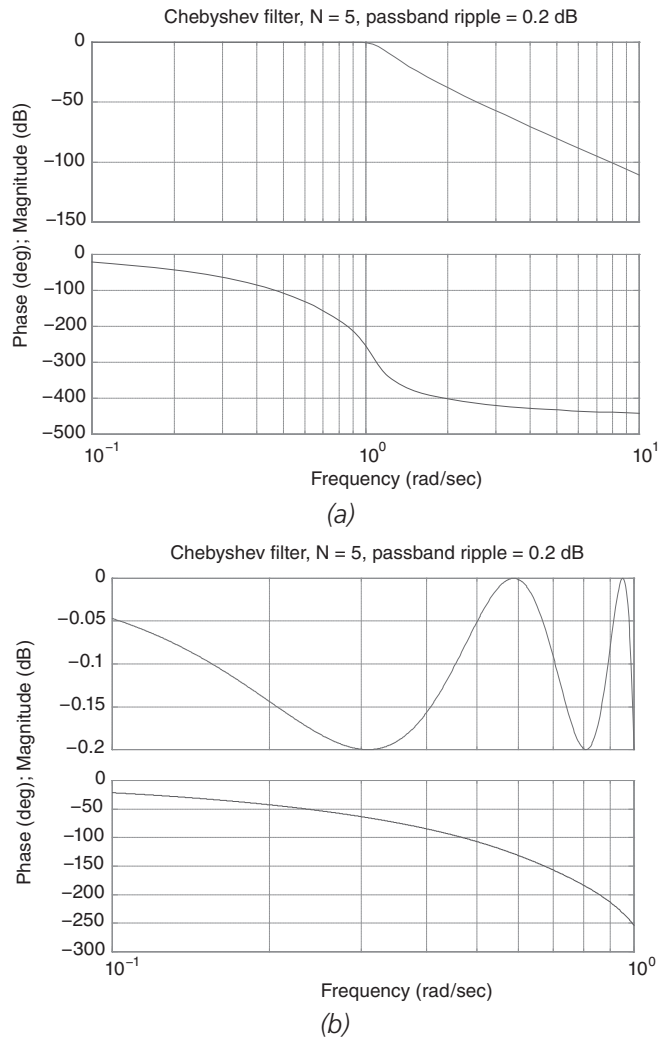


Figure 14-6: Chebyshev filter frequency response, order $N = 5$, passband ripple 0.20dB. (a) Overall response. (b) Detail of passband ripple. Note that the -3 dB cutoff frequency is slightly higher than 1 radian/second. The Chebyshev approximation shown here results in a passband ripple frequency of 1 radian/second.

Chapter 14

A tabulation of transfer functions and -3dB frequencies for three different types of Chebyshev filters are shown in **Table 14-2**, **Table 14-3** and **Table 14-4**.

Table 14-2: Transfer functions⁴ for Chebyshev filters of varying filter order N , with filter ripple frequency 1 rad/sec. , passband ripple $= 0.1\text{ dB}$. The -3dB point ω_c is somewhat higher, as shown.

N	Transfer function	$\omega_{3\text{dB}}$ (rad/s)
2	$\frac{3.314}{s^2 + 2.3724s + 3.314}$	1.94
3	$\frac{1.6381}{(s + 0.9694)(s^2 + 0.9694s + 1.6899)}$	1.38
4	$\frac{0.8285}{(s^2 + 0.5283s + 1.33)(s^2 + 1.2755s + 0.6229)}$	1.21
5	$\frac{0.4095}{(s + 0.5389)(s^2 + 0.3331s + 1.1949)(s^2 + 0.8720s + 0.6359)}$	1.13
6	$\frac{0.2071}{(s^2 + 0.2294s + 1.1294)(s^2 + 0.6267s + 0.6964)(s^2 + 0.8561s + 0.2634)}$	1.09
7	$\frac{0.1024}{(s + 0.3768)(s^2 + 0.1677s + 1.0924)(s^2 + 0.4698s + 0.7532)(s^2 + 0.6789s + 0.3302)}$	1.07
8	$\frac{0.0518}{(s^2 + 0.12805s + 1.0695)(s^2 + 0.3644s + 0.7989)(s^2 + 0.5454s + 0.4162)(s^2 + 0.6433s + 0.1456)}$	1.05

Table 14-3: Transfer functions for Chebyshev filters of varying filter order N , with filter ripple frequency 1 rad/sec. , passband ripple $= 0.20\text{ dB}$.

N	Transfer function	$\omega_{3\text{dB}}$ (rad/s)
2	$\frac{2.3568}{s^2 + 1.9271s + 2.3568}$	1.67
3	$\frac{1.1516}{(s + 0.8146)(s^2 + 0.8146s + 1.4136)}$	1.28

⁴ These transfer functions were generated using MATLAB and the CHEBY1 function that calculates the pole locations of Chebyshev filters with a given passband ripple.

N	Transfer function	ω_{3dB} (rad/s)
4	$\frac{0.5892}{(s^2 + 0.4496s + 1.1987)(s^2 + 1.0855s + 0.4916)}$	1.16
5	$\frac{0.2879}{(s + 0.4614)(s^2 + 0.2852s + 1.1174)(s^2 + 0.7466s + 0.5584)}$	1.10
6	$\frac{0.1473}{(s^2 + 0.1970s + 1.0779)(s^2 + 0.5383s + 0.6449)(s^2 + 0.7354s + 0.2119)}$	1.07
7	$\frac{0.0720}{(s + 0.3243)(s^2 + 0.1433s + 1.0557)(s^2 + 0.4044s + 0.7164)(s^2 + 0.5844s + 0.2934)}$	1.05
8	$\frac{0.0368}{(s^2 + 0.1103s + 1.0418)(s^2 + 0.3141s + 0.7712)(s^2 + 0.4700s + 0.3886)(s^2 + 0.5544s + 0.1180)}$	1.04

Table 14-4: Transfer functions for Chebyshev filters of varying filter order N, with filter ripple frequency = 1 rad/sec., passband ripple = 0.50dB.

N	Transfer function	ω_{3dB} (rad/s)
2	$\frac{1.5162}{s^2 + 1.4256s + 1.5162}$	1.39
3	$\frac{0.7157}{(s + 0.6265)(s^2 + 0.6265s + 1.1424)}$	1.17
4	$\frac{0.3791}{(s^2 + 0.3507s + 1.0635)(s^2 + 0.8467s + 0.3564)}$	1.09
5	$\frac{0.1789}{(s + 0.3623)(s^2 + 0.2239s + 1.0358)(s^2 + 0.5762s + 0.4768)}$	1.06
6	$\frac{0.0948}{(s^2 + 0.1553s + 1.0230)(s^2 + 0.4243s + 0.5900)(s^2 + 0.5796s + 0.1570)}$	1.04
7	$\frac{0.0447}{(s + 0.2562)(s^2 + 0.1140s + 1.0161)(s^2 + 0.3194s + 0.6769)(s^2 + 0.4616s + 0.2539)}$	1.03
8	$\frac{0.0237}{(s^2 + 0.0872s + 1.0119)(s^2 + 0.2484s + 0.7413)(s^2 + 0.3718s + 0.3587)(s^2 + 0.4386s + 0.0881)}$	1.02

A comparison of group delay responses for Chebyshev filters⁵ of varying order are shown in **Figure 14-7**. Note the peaking in the group delay near the cutoff frequency. This peaking, and group delay ripple in the passband, become more pronounced as we increase the passband ripple of the Chebyshev filter.

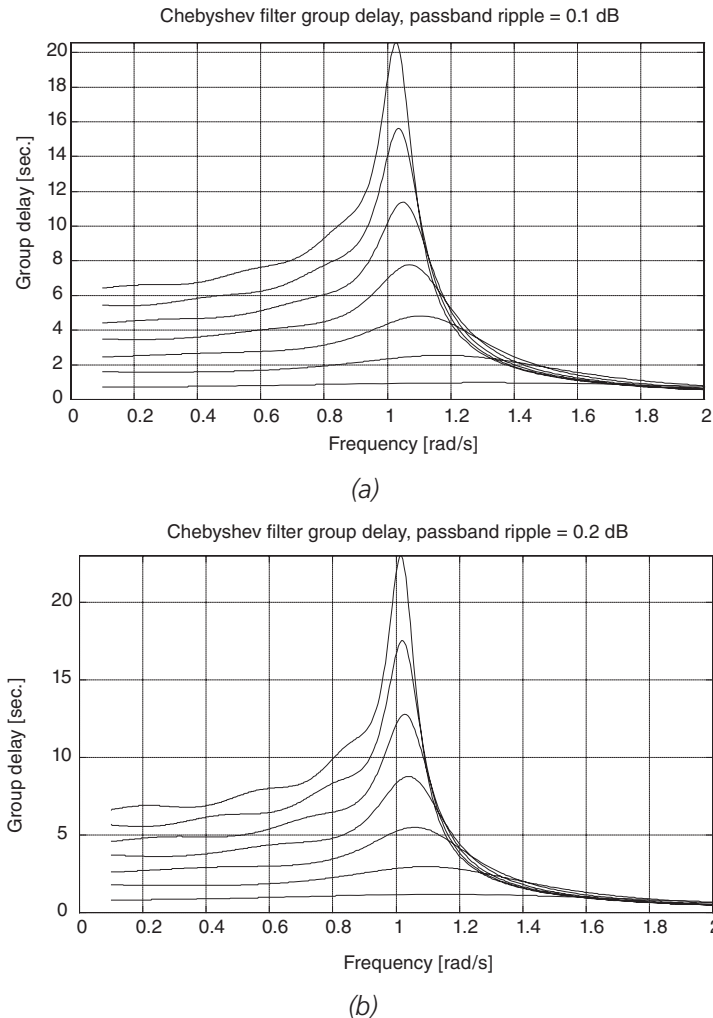


Figure 14-7: Comparison of group delay responses for Chebyshev filters with passband ripple 0.1dB, 0.2dB and 0.5dB and cutoff frequencies near 1 radian/sec, and filter orders $N = 2$ to 8.
(Continued on following page.)

⁵ In each case, the filters are normalized for a frequency at which the ripple band is first exceeded of 1 radian/second. Therefore, the cutoff frequency for each filter is approximately 1 radian/sec., but not exactly.

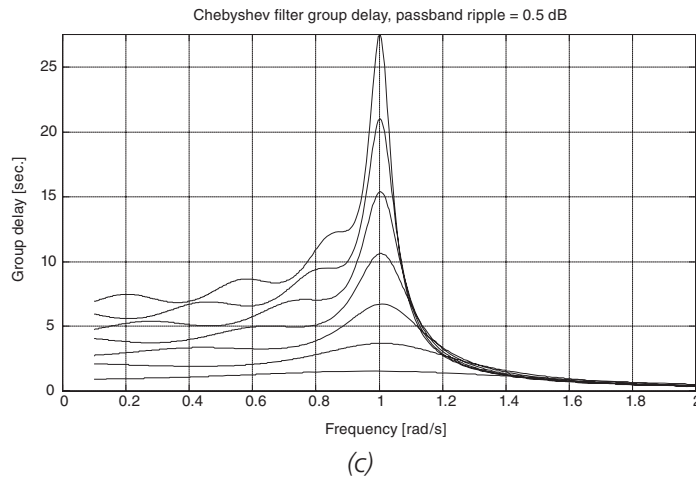


Figure 14-7 (continued): Comparison of group delay responses for Chebyshev filters with passband ripple 0.1dB, 0.2dB and 0.5dB and cutoff frequencies near 1 radian/sec.

Bessel Filter

The Bessel filter is optimized to provide constant group delay in the filter passband, while sacrificing sharpness in the frequency response. We note these effects with reference to **Figure 14-8**, **Figure 14-9** and **Figure 14-10**. The frequency response is gradual in the transition band, and there is little or no overshoot in the step response. The group delay exhibits very flat response in the passband.

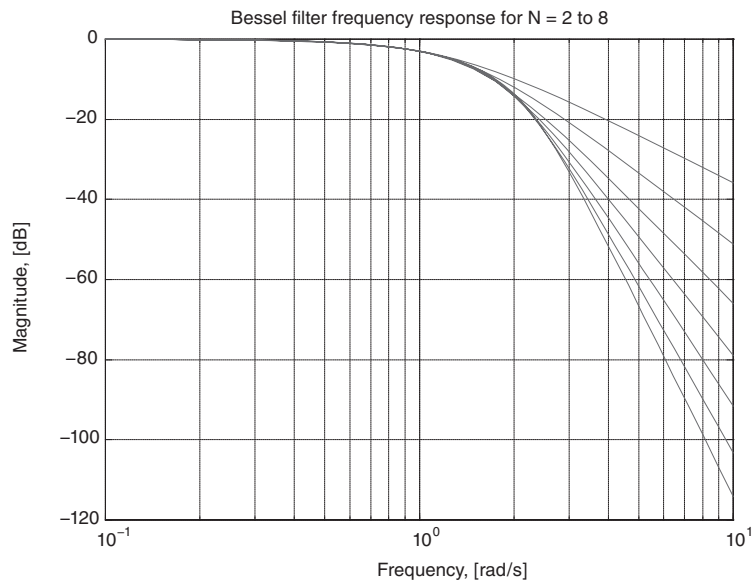


Figure 14-8: Bessel filter frequency response, order N = 2 to 8, shown for filters with -3dB cutoff frequency of 1 radian/second.

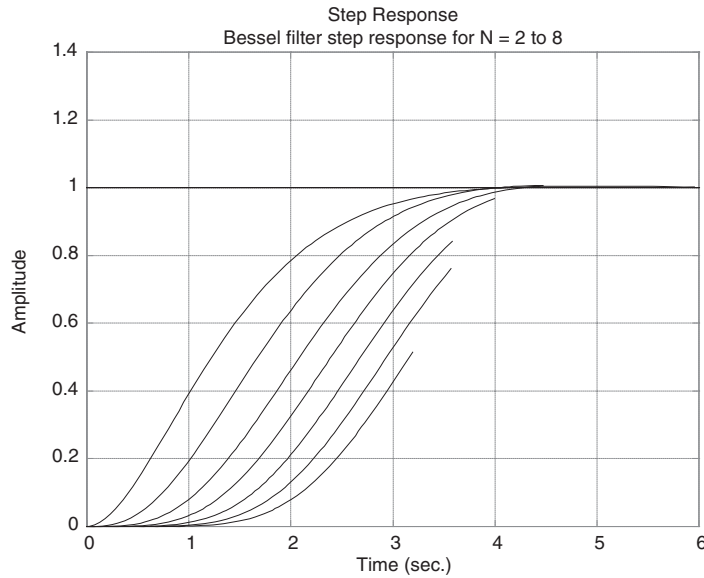


Figure 14-9: Bessel filter step response, order $N = 2$ to 8, shown for filters with -3dB cutoff frequency of 1 radian/second.

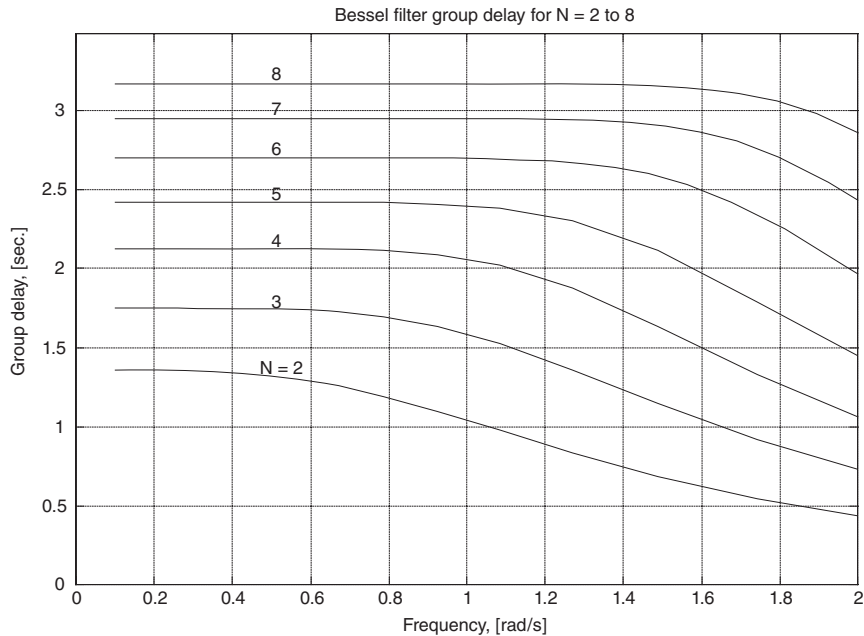


Figure 14-10: Bessel filter group delay response, order $N = 2$ to 8, shown for filters with -3dB cutoff frequency of 1 radian/second. Note that as the filter order increases, the flatness of the group delay response in the passband improves.

Table 14-5: Pole locations for Bessel filters of varying filter order N, with filter cutoff frequency $\omega_c = 1$ rad/sec.

N	Real Part ($-\sigma$)	Imaginary Part ($\pm j\omega$)
2	1.1030	0.6368
3	1.0509 1.3270	1.0025
4	1.3596 0.9877	0.4071 1.2476
5	1.3851 0.9606 1.5069	0.7201 1.4756
6	1.5735 1.3836 0.9318	0.3213 0.9727 1.6640
7	1.6130 1.3797 0.9104 1.6853	0.5896 1.1923 1.8375
8	1.7627 0.8955 1.3780 1.6419	0.2737 2.0044 1.3926 0.8253

The resultant transfer functions of the Bessel filters are shown in **Table 14-6**.

Table 14-6: Transfer function broken up into first-order and second-order quadratic factors for Bessel filters of varying filter order N, with filter cutoff frequency $\omega_c = 1$ rad/sec.

N	Transfer function
2	$\frac{1.6221}{s^2 + 2.206s + 1.6221}$
3	$\frac{2.7992}{(s + 1.3270)(s^2 + 2.1018s + 2.1094)}$
4	$\frac{5.1002}{(s^2 + 2.7192s + 2.0142)(s^2 + 1.9754s + 2.5321)}$
5	$\frac{11.3845}{(s + 1.5069)(s^2 + 2.7702s + 2.4370)(s^2 + 1.9212s + 3.1001)}$
6	$\frac{26.8328}{(s^2 + 3.1470s + 2.5791)(s^2 + 2.7672s + 2.8605)(s^2 + 1.8636s + 3.6371)}$

N	Transfer function
7	$\frac{69.5099}{(s + 1.6853)(s^2 + 3.2262s + 2.9497)(s^2 + 2.7594s + 3.3251)(s^2 + 1.8208s + 4.2052)}$
8	$\frac{198.7746}{(s^2 + 3.5254s + 3.1820)(s^2 + 1.7910s + 4.1895)(s^2 + 2.7560s + 3.8382)(s^2 + 3.2838s + 3.3770)}$

Comparison of Responses of Different Filter Types

A comparison of frequency responses for three different types of an $N = 5$ filter is shown in **Figure 14-11**. Note that, as expected, the Chebyshev filter has the sharpest cutoff characteristics, while the Bessel response is more gradual in the stopband.

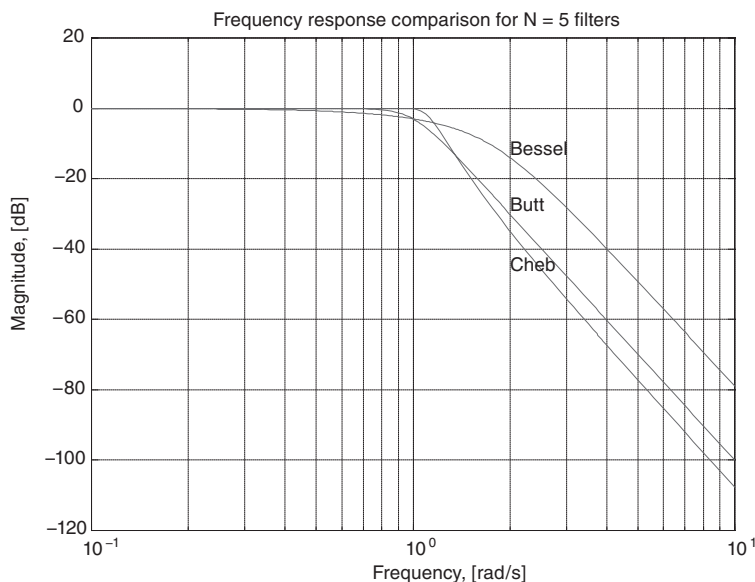


Figure 14-11: Comparison of frequency response for $N = 5$ Butterworth, $N = 5$ Chebyshev with 0.1dB ripple and $N = 5$ Bessel filter, each with a cutoff frequency of 1 radian/second.

A comparison of step responses for the three different filters is shown in **Figure 14-12**. Note that the Bessel has minimal overshoot, while the overshoots of the Butterworth and Chebyshev filters are comparable. A Chebyshev filter with more passband ripple would have higher overshoot.

In **Figure 14-13** we compare the group delay responses of the three filters. We note that the Bessel filter has flat group delay response in the passband, while the Butterworth has moderate variation in group delay in the passband. The Chebyshev exhibits pronounced group delay peaking in the passband.

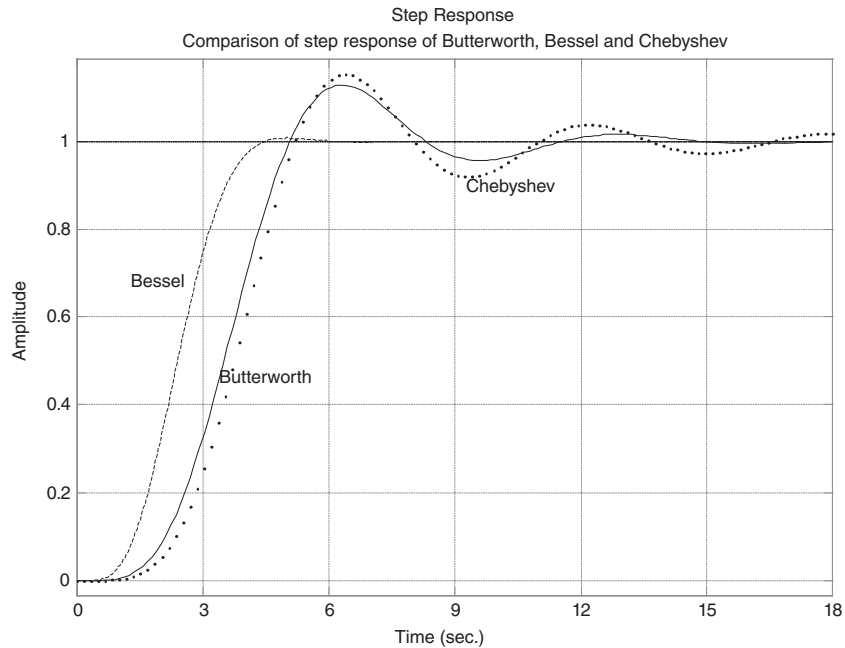


Figure 14-12: Comparison of step response for $N = 5$ Butterworth, Chebyshev (0.1dB ripple) and Bessel filter, each with a cutoff frequency of 1 radian/second. Legend: Butterworth (solid), Bessel (dashed), Chebyshev (dotted). Note that there is minimal overshoot in the Bessel response, while the Butterworth and Chebyshev responses are comparable.

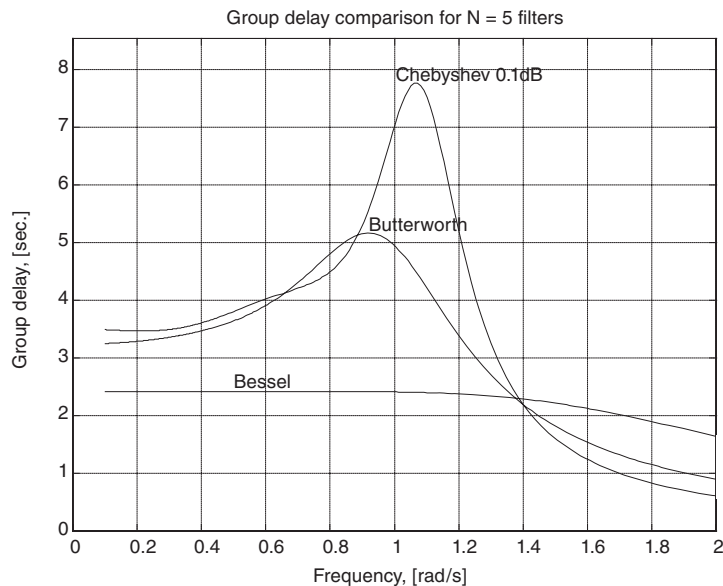


Figure 14-13: Comparison of group delay response for $N = 5$ Butterworth, Chebyshev (0.1dB ripple) and Bessel filter, each with a cutoff frequency of 1 radian/second.

Filter Implementation

So, you've done all this hard work to figure out the filter type and order, but how do you build it in practice? There are several methods and we'll discuss a few of them in this next section.

Ladder

For high-frequency filters, one option is to build a passive ladder using resistors, inductors and capacitors. The topology for an N th-order ladder filter suitable for implementing Butterworth, Chebyshev and Bessel filters is shown in **Figure 14-14a**. The filter is made with alternating inductors and capacitors, with source and termination resistors. We note that, for an even-order filter, the filter terminates with a resistor and capacitor and for an odd-order filter the filter terminates with an inductor and resistor. Ladders for $N = 4$ and $N = 5$ filters are shown in **Figure 14-14b** and **Figure 14-14c**, respectively.

Values have been extensively tabulated for filters with cutoff frequency 1 radian per second. The termination resistor for this normalized filter is 1.0 ohms, and the filter can have any input resistance R_s that you desire. Tabulated ladder element values for Butterworth, Bessel and Chebyshev filters of various orders are shown in **Table 14-7** through **Table 14-11**.

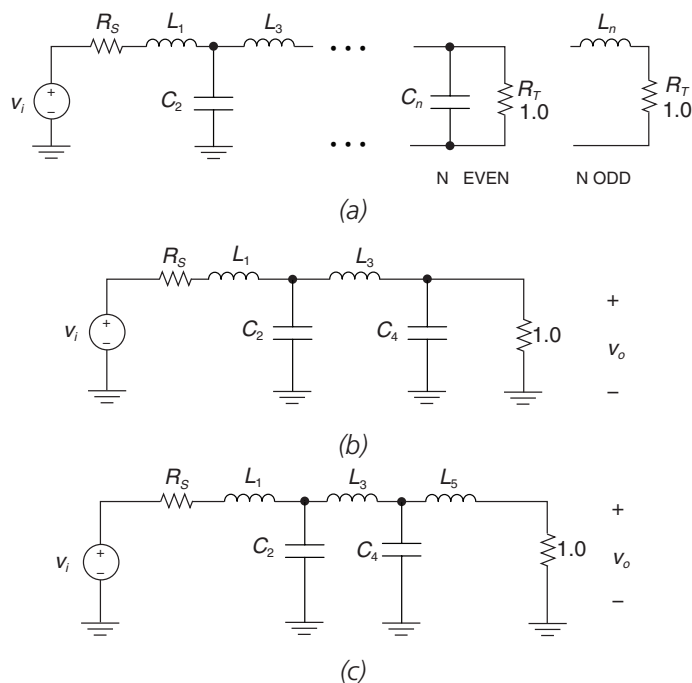


Figure 14-14: Ladder filter of order N , suitable for implementing Butterworth, Chebyshev and Bessel filters. (a) Generic filter. (b) $N = 4$ ladder. (c) $N = 5$ ladder.

Table 14-7: Butterworth inductor⁶ and capacitor values of varying filter order N, with filter cutoff frequency $\omega_c = 1$ rad/sec., and termination resistance $R_T = 1.0\Omega$. Note that for this case, all values of the source resistance are $R_s = 1.0\Omega$.

N	R_s	L_1	C_2	L_3	C_4	L_5	C_6	L_7	C_8
2	1.0	1.4142	1.4142						
3	1.0	1.0000	2.0000	1.0000					
4	1.0	0.7654	1.8478	1.8478	0.7654				
5	1.0	0.6180	1.6180	2.0000	1.6180	0.6180			
6	1.0	0.5176	1.4142	1.9319	1.9319	1.4142	0.5176		
7	1.0	0.4450	1.2470	1.8019	2.0000	1.8019	1.2470	0.4450	
8	1.0	0.3902	1.1111	1.6629	1.9616	1.9616	1.6629	1.1111	0.3902

Table 14-8: Bessel inductor and capacitor values of varying filter order N, with filter cutoff frequency $\omega_c = 1$ rad/sec., and termination resistance $R_T = 1.0\Omega$. Note that for this case, all values of the source resistance are $R_s = 1.0\Omega$.

N	R_s	L_1	C_2	L_3	C_4	L_5	C_6	L_7	C_8
2	1.0	0.5755	2.1478						
3	1.0	0.3374	0.9705	2.2034					
4	1.0	0.2334	0.6725	1.0815	2.2404				
5	1.0	0.1743	0.5072	0.8040	1.1110	2.2582			
6	1.0	0.1365	0.4002	0.6392	0.8538	1.1126	2.2645		
7	1.0	0.1106	0.3259	0.5249	0.7020	0.8690	1.1052	2.2659	
8	1.0	0.0919	0.2719	0.4409	0.5936	0.7303	0.8695	1.0956	2.2656

Table 14-9: 0.1dB Chebyshev filter inductor and capacitor values, with filter cutoff frequency $\omega_c = 1$ rad/sec., and termination resistance $R_T = 1.0\Omega$. Note that the values of the source resistance vary from filter order to filter order.

N	R_s	L_1	C_2	L_3	C_4	L_5	C_6	L_7	C_8
2	1.3554	1.2087	1.6382						
3	1.0	1.4328	1.5937	1.4328					
4	1.3554	0.9924	2.1476	1.5845	1.3451				
5	1.0	1.3013	1.5559	2.2411	1.5559	1.3013			
6	1.3554	0.9419	2.0797	1.6581	2.2473	1.5344	1.2767		
7	1.0	1.2615	1.5196	2.2392	1.6804	2.2392	1.5196	1.2615	
8	1.3554	0.9234	2.0454	1.6453	2.2826	1.6841	2.2300	1.5091	1.2515

⁶ All inductor and capacitor value charts adopted from A. Zverev, *Handbook of Filter Synthesis*, John Wiley, 1967. For other source resistances R_s , the reader is invited to visit this reference.

Table 14-10: 0.25dB Chebyshev filter inductor and capacitor values, with filter cutoff frequency $\omega_c = 1$ rad/sec., and termination resistance $R_T = 1.0\Omega$.

N	R_s	L_1	C_2	L_3	C_4	L_5	C_6	L_7	C_8
2	2.0	0.6552	2.7632						
3	1.0	1.6325	1.4360	1.6325					
4	2.0	0.6747	3.6860	1.0247	1.8806				
5	1.0	1.5046	1.4436	2.4050	1.4436	1.5046			
6	2.0	0.6867	3.2074	0.9308	3.8102	1.2163	1.7088		
7	1.0	1.5120	1.4169	2.4535	1.5350	2.4535	1.4169	1.5120	

Table 14-11: 0.5dB Chebyshev filter inductor and capacitor values, with filter cutoff frequency $\omega_c = 1$ rad/sec., and termination resistance $R_T = 1.0\Omega$.

N	R_s	L_1	C_2	L_3	C_4	L_5	C_6	L_7	C_8
2	1.9841	0.9827	1.9497						
3	1.0	1.8636	1.2804	1.8636					
4	1.9841	0.9202	2.5864	1.3036	1.8258				
5	1.0	1.8068	1.3025	2.6912	1.3025	1.8068			
6	1.9841	0.9053	2.5774	1.3675	2.7133	1.2991	1.7961		
7	1.0	1.7896	1.2961	2.7177	1.3848	2.7177	1.2961	1.7896	
8	1.9841	0.8998	2.5670	1.3697	2.7585	1.3903	2.7175	1.2938	1.7852

Example 14.1: Design example: Fifth-order Chebyshev filter with 0.5dB passband ripple

We'll use the filter charts to design a fifth-order 1-MHz low-pass Chebyshev filter with 0.5dB ripple in the passband. From the filter chart (**Table 14-11**) we find the corresponding values for a filter with cutoff frequency 1 radian/second as:

$$R_s = R_T = 1\Omega$$

$$L_1 = 1.8068$$

$$C_2 = 1.3025$$

$$L_3 = 2.6914$$

$$C_4 = 1.3025$$

$$L_5 = 1.8068$$

We next need to pick more reasonable values for source and termination resistors (instead of the 1.0Ω normalized values). For this design example we'll choose $R_s = R_T = 50\Omega$. We now make use of an unnormalization process to transform the filter with cutoff of 1 radian per second to our desired frequency of 1 MHz. The unnormalization process is:

$$C = \frac{C_n}{2\pi f_c R}$$

$$L = \frac{L_n R}{2\pi f_c}$$
[14-6]

where C_n and L_n are the normalized values found from the filter charts, f_c is the desired new cutoff frequency, and R is the resistor value used in the new filter. Applying this process to our filter results in:

$$\begin{aligned} L'_1 &= \frac{L_1 R}{2\pi f_c} = \frac{(1.8068)(50)}{(2\pi)(10^6)} = 14.378 \mu\text{H} \\ C'_2 &= \frac{C_2}{2\pi f_c R} = \frac{(1.3025)}{(2\pi)(10^6)(50)} = 4146 \text{ pF} \\ L'_3 &= \frac{L_3 R}{2\pi f_c} = \frac{(2.6914)(50)}{(2\pi)(10^6)} = 23.407 \mu\text{H} \\ C'_4 &= \frac{C_4}{2\pi f_c R} = \frac{(1.3025)}{(2\pi)(10^6)(50)} = 4146 \text{ pF} \\ L'_4 &= \frac{L_4 R}{2\pi f_c} = \frac{(1.8068)(50)}{(2\pi)(10^6)} = 14.378 \mu\text{H} \end{aligned} \quad [14-7]$$

The resultant circuit and frequency response for the 1-MHz filter is shown in **Figure 14-15**.

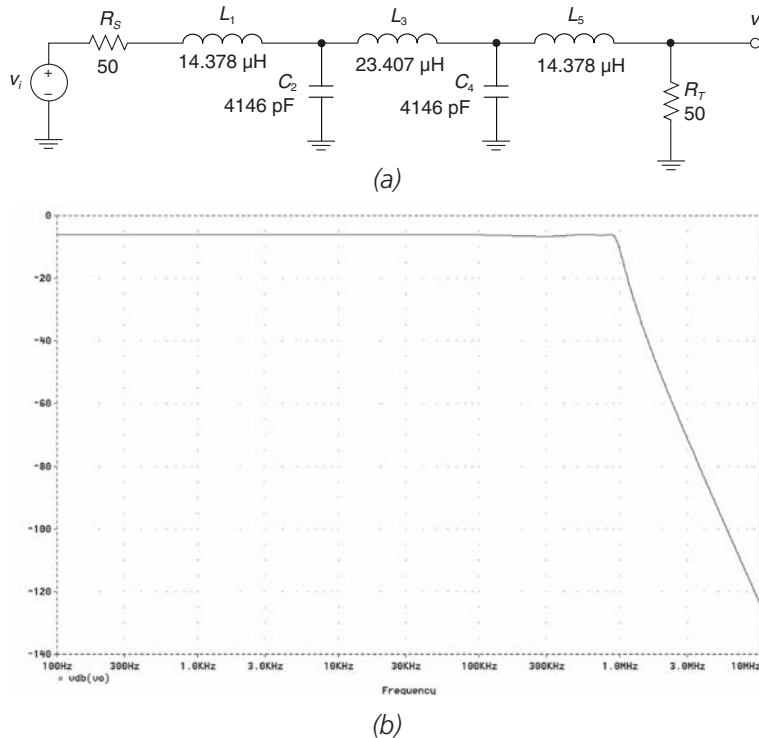


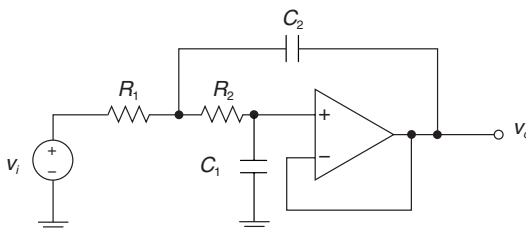
Figure 14-15: 1-MHz, 0.5dB Chebyshev low-pass filter implemented as a ladder. (a) PSPICE circuit. (b) Frequency response. Note that the DC gain of this filter is -6dB (or a factor of $1/2$) due to the R_S - R_T resistive divider.

Filter Implementation—Active

We can also directly implement the filter transfer functions using active filters, such as the Sallen-Key filter. The Sallen-Key filter (**Figure 14-16**) generates the transfer function

$$\frac{v_o}{v_i} = \frac{1}{R_1 R_2 C_1 C_2 s^2 + [R_2 C_2 + R_1 C_2]s + 1} \quad [14-8]$$

Figure 14-16: Sallen-Key circuit (with DC gain = 1).

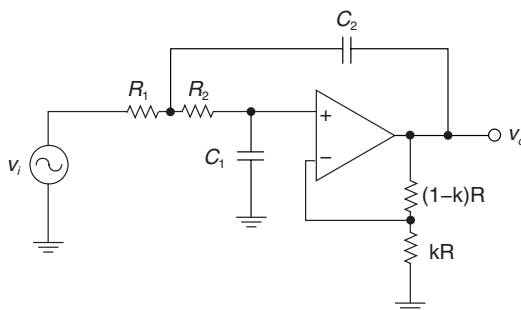


For instance, to implement a fourth-order filter, you could cascade two Sallen-Key circuits, provided you know where all the filter poles are.

A variation on the theme is the Sallen-Key circuit with adjustable DC gain, which also adjusts the damping of the filter, as shown in **Figure 14-17**. The transfer function of this filter is:

$$\frac{v_{out}}{v_{in}} = \left(\frac{1}{k}\right) \frac{1}{R_1 R_2 C_1 C_2 s^2 + \left[R_2 C_2 + R_1 C_2 + R_1 C_1 \left(1 - \frac{1}{k}\right)\right]s + 1} \quad [14-9]$$

Figure 14-17: Sallen-Key Circuit (with adjustable Q). The adjustment can be made by implementing the $(1-k)R$ and kR resistor with a potentiometer.



Some comments on elliptic (or “brick wall”) filters

The detailed design of elliptic filters is beyond the scope of this book, but a few comments are in order. Elliptic filters, also called “brick wall” filters, have very sharp filter cutoff characteristics. Again, this is done at the expense of very nonlinear group delay. One flavor of elliptic filter has zero ripple in the passband but finite ripple in the stopband. This is accomplished by having zeroes in the transfer function. A 1 radian/second elliptic low-pass filter is shown in **Figure 14-18**. Note that the filter has parallel LC sections. These sections generate zeros in the transfer function.

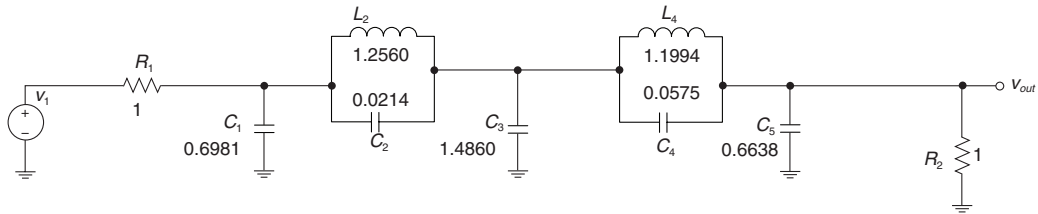
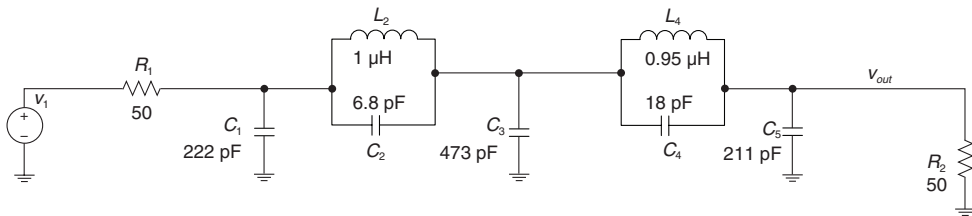
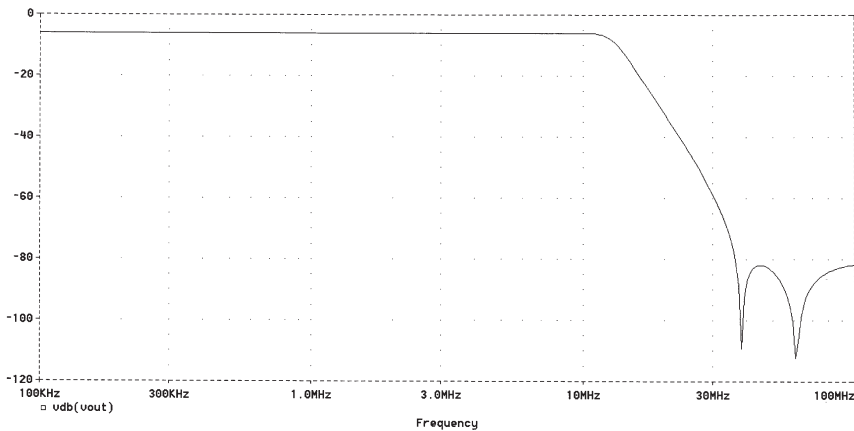


Figure 14-18: Elliptic low-pass filter prototype (1 r/s cutoff).

We can now unnormalize this filter as before to generate a filter with a cutoff frequency of 10 MHz (**Figure 14-19a**). The frequency response (**Figure 14-19b**) shows a very fast rolloff in the transition band, and has a minimum attenuation floor.



(a)



(b)

Figure 14-19: Unnormalized elliptic low-pass filter prototype (10-MHz cutoff).
(a) PSPICE Circuit. (b) Frequency response.

Example 14.2: Design Example: 40-Hz Sallen-Key with Adjustable Q

Following is a design of a 40-Hz low-pass filter with adjustable Q:

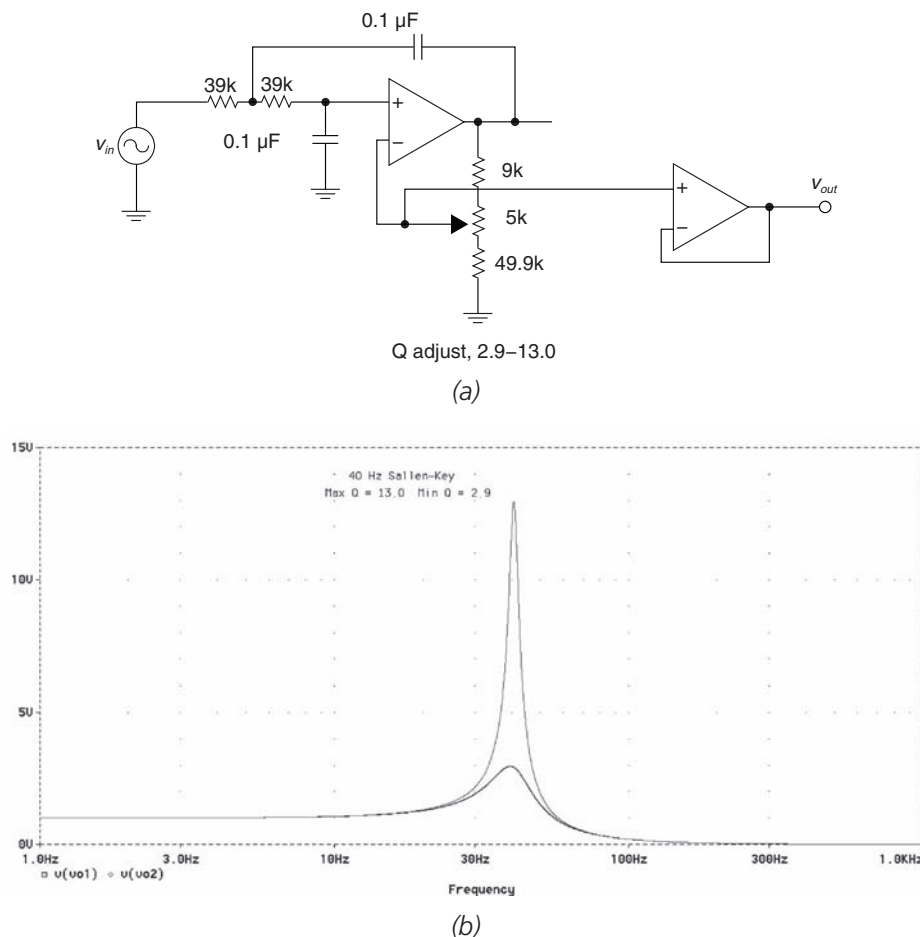
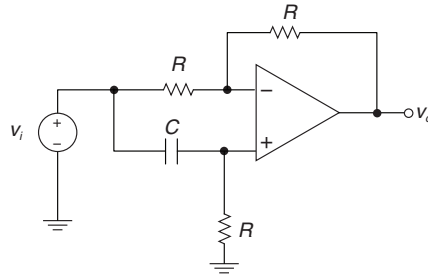


Figure 14-20: 40 Hz Sallen-Key filter with adjustable Q. (a) Circuit. (b) SPICE simulation.

All-pass filters

An all-pass filter is a filter that has a magnitude response of unity, but which provides phase shift. You can use all-pass filters to tailor group delay responses in your filters. You may find that you will need to cascade your filter with an all-pass filter in order to meet the group delay specification. A first-order all-pass circuit is shown in **Figure 14-21**. Note that this all-pass provides a DC gain of -1 . If you want, you can cascade an inverting op-amp stage with the all-pass to take care of this phase inversion.

Figure 14-21: First-order all-pass filter.



The transfer function, angle and group delay for a first-order all-pass filter are:

$$H(s) = \frac{RCs - 1}{RCs + 1} = \frac{s - \frac{1}{RC}}{s + \frac{1}{RC}} = \frac{s - a}{s + a} \quad [14-10]$$

$$\angle H(s) = -2 \tan^{-1} \frac{\omega}{a}$$

$$D(j\omega) = \frac{2a}{a^2 + \omega^2}$$

The group delay characteristics for $a = 1$ r/s is shown in **Figure 14-22**. Note that the DC delay is twice the value of RC .

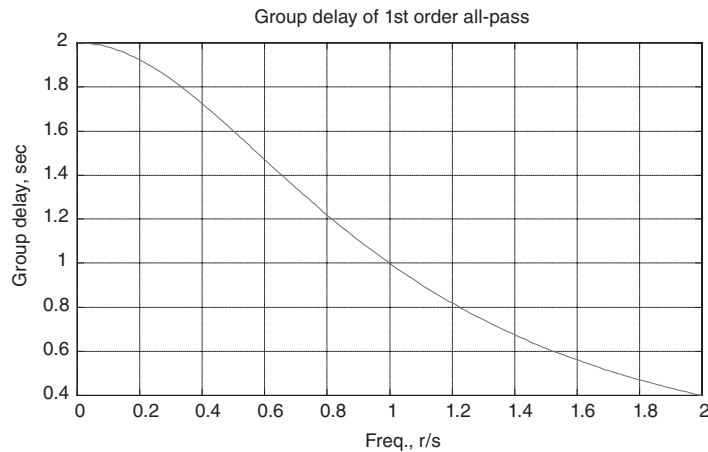


Figure 14-22: First-order all-pass filter group delay response.

Example 14.3: Design case study: 1-MHz low-pass filter

In this case study, we'll design and simulate an analog low-pass filter meeting the following specifications (**Table 14-12**):

Table 14-12: Design case study specification.

Filter type	Low-pass
Nominal -3dB bandwidth	1 MHz
Passband gain	0dB nominal, within 0.25dB of nominal up to 750 kHz
Attenuation	> 50dB at 2.5 MHz
Group delay response	Group delay variation from the DC value less than 1000 nanoseconds up to 1 MHz

This design could be implemented with either a ladder filter with passive elements (resistors, inductors and capacitor) or a cascade of Sallen-Key active second-order sections. If you attempt to use a Bessel filter, the filter order will be quite high. If you look at elliptic filters the group delay specification may be very difficult to meet.

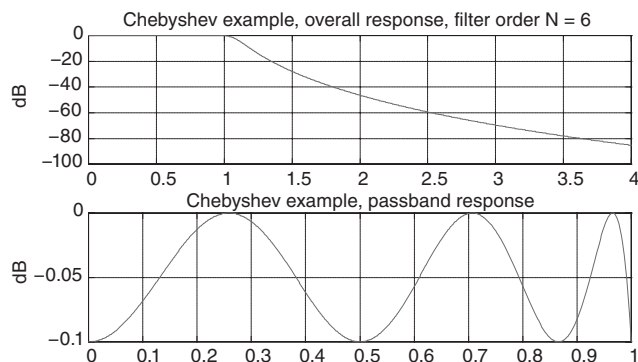
In this design $\omega_s/\omega_c = 2.5$ and minimum attenuation at ω_s is -50dB. A Bessel filter was not implemented due to the high filter order required. In order to meet the specification for -50dB gain at a normalized frequency of 2.5, a Bessel filter of order $N > 10$ would be needed. A sixth-order Butterworth filter would *almost* make the specification, since

$$20 \log_{10} \left(\frac{1}{\sqrt{1 + 2.5^{12}}} \right) = -47.8 \text{ dB}$$

So, if we use a Butterworth, we'll need at least $N = 7$. The advantage of the Butterworth is that there is no magnitude ripple in the passband.

Another alternative will be to try a Chebyshev design with $N = 7$ or less and some passband ripple. From our previous work on the Chebyshev filter, it looks like an $N = 6$ Chebyshev with 0.1dB ripple in the passband will meet the specification. An $N = 5$ Chebyshev with 0.25dB passband ripple barely misses the attenuation spec. So, let's go with the $N = 6$, 0.1dB Chebyshev design (**Figure 14-23**).

Figure 14-23: Chebyshev design, $N = 6$, 0.1 dB passband ripple.



Using a ladder filter topology for an $N = 6$ Chebyshev with 0.1dB ripple, normalized component values for $\omega_c = 1$ r/s and $R_T = 1\Omega$ are:

$$R_s = 1.3554$$

$$L_1 = 0.9419$$

$$C_2 = 2.0797$$

$$L_3 = 1.6581$$

$$C_4 = 2.2473$$

$$L_5 = 1.5344$$

$$C_6 = 1.2767$$

Remember, we scale the normalized filter by:

$$L = \frac{L_{norm} R}{\omega_c}$$

$$C = \frac{C_{norm}}{\omega_c R}$$

From this, we calculate the unnormalized component values for a filter with -3dB point of 1 MHz (**Figure 14-24**), and using a termination resistor value of 75Ω . The overall magnitude of the frequency response of the Chebyshev ladder filter is shown in **Figure 14-25**.

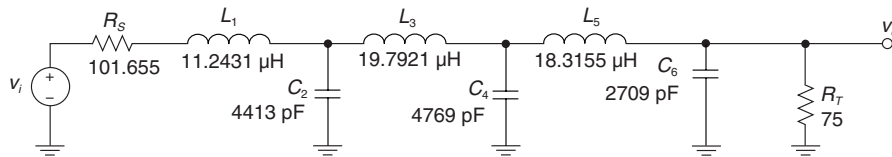


Figure 14-24: Chebyshev ladder filter design.

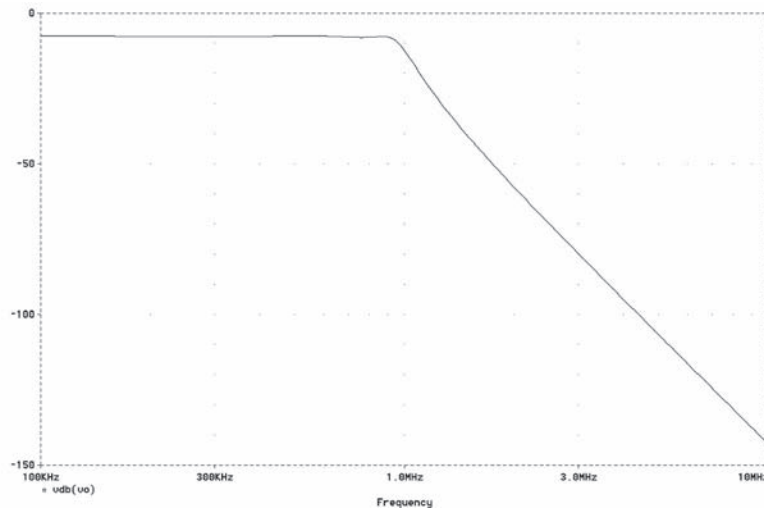


Figure 14-25: Chebyshev ladder filter magnitude of the frequency response (from SPICE). Vertical scale in dB.

The stopband detail shows that the magnitude is attenuated $> 60\text{dB}$ at 2.5 MHz , as expected (**Figure 14-26**).

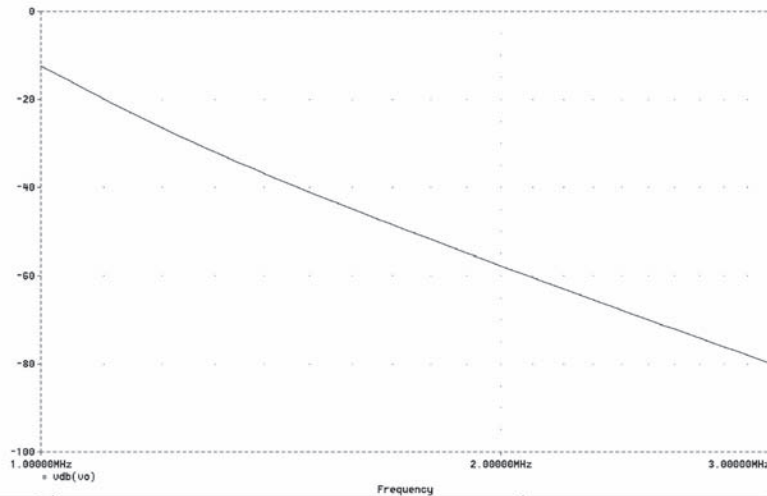


Figure 14-26: Chebyshev ladder filter design frequency response, stopband detail. Vertical scale in dB.

Note that there is slightly more than expected ripple in the passband. This may be due to roundoff error in the ladder components (**Figure 14-27**).

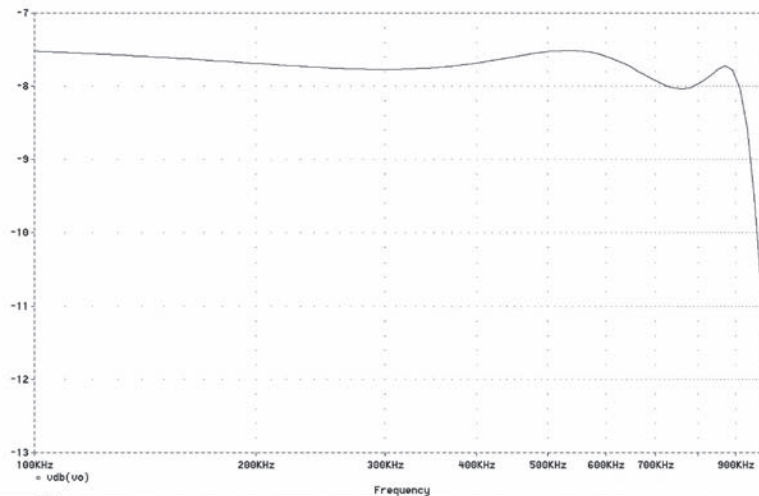


Figure 14-27: Chebyshev ladder filter design frequency response, passband detail. Vertical scale in dB.

SPICE results show that the variation in group delay in the DC to 1-MHz range is approximately 1200 nanoseconds, which violates the group delay specification (**Figure 14-28**).

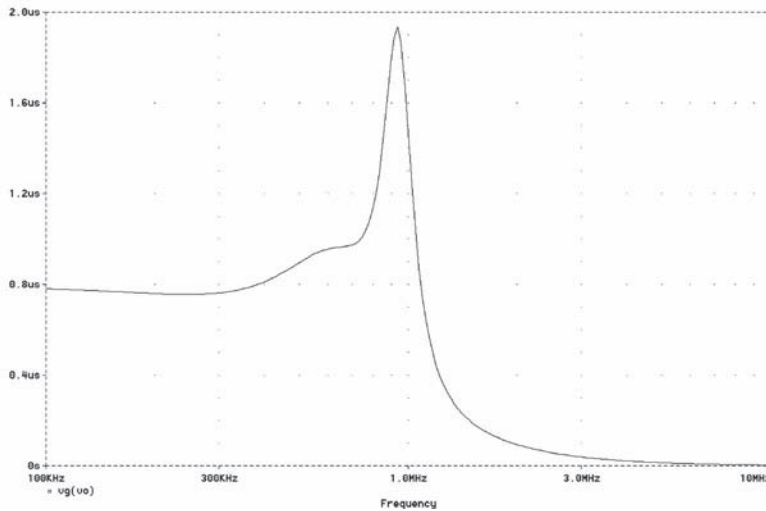
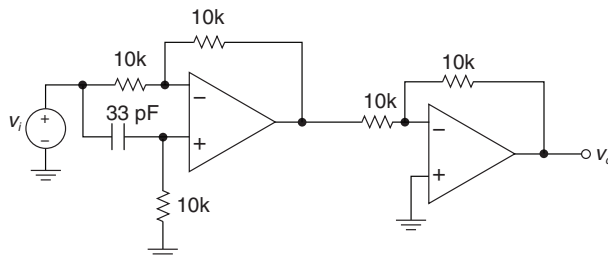


Figure 14-28: Chebyshev ladder filter design group delay. There is approximately 1200 nanoseconds of group delay variation in the passband.

From the above, we see that the delay variation in the DC to 1-MHz range is over 1000 nanoseconds. So, let's cascade a first-order all-pass network in an attempt to fill in the group delay hole below 1 MHz (**Figure 14-29**). We'll assume that we have ideal op-amps at our disposal. What this means is that we need to choose op-amps with gain-bandwidth product much higher than our frequency range of interest, or greater than 1 MHz. The low-frequency delay of a first-order all-pass is $2RC$, so in this case we've chosen a DC delay of 660 nanoseconds.

Figure 14-29: Group delay equalizer for filter example.



The delay-equalized filter meets the group delay specification. The peak-to-peak delay variation in the passband is approximately 900 nanoseconds (**Figure 14-30**), and hence we meet the group delay specification.

The step response of original and delay equalized circuits (**Figure 14-31**) shows that the equalized circuit has less overshoot in the step response, as expected. However, the delay through the filter is increased (due to the all-pass network).

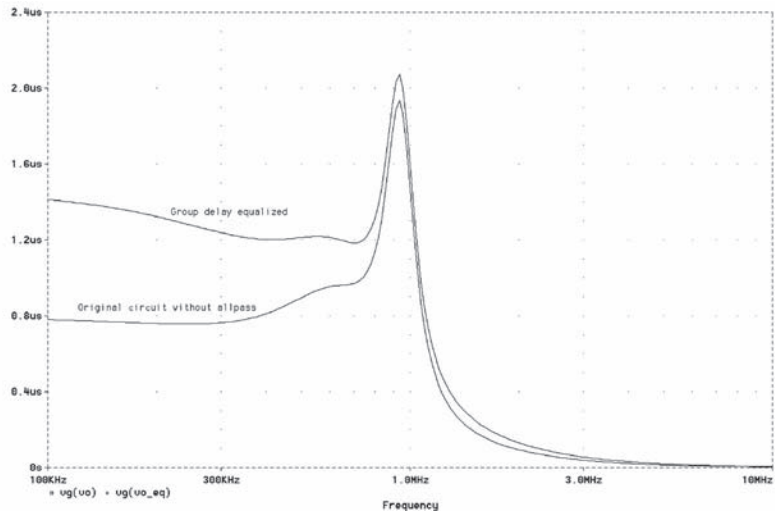


Figure 14-30: Chebyshev ladder filter design group delay equalized, compared to original circuit.

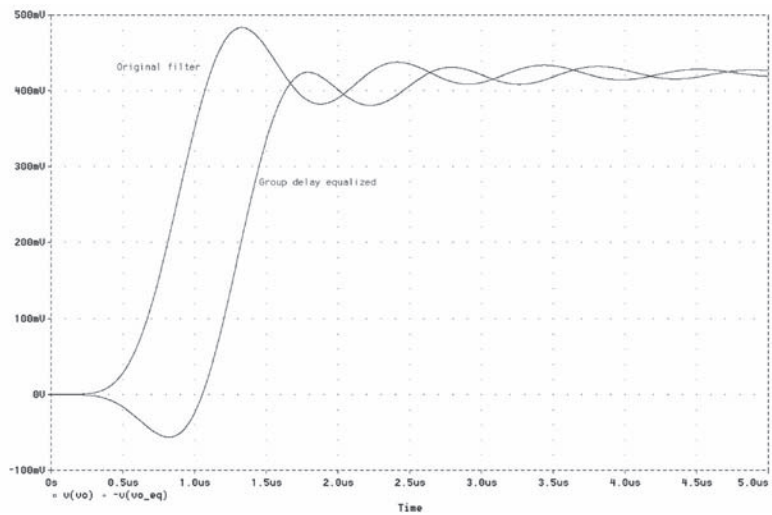


Figure 14-31: Chebyshev ladder filter design group delay equalized, step response.

Example 14.4: Alternate design using Butterworth filter

Here's a seventh-order Butterworth design, which also meets the group delay specification *without* any further all-pass filtering. In this case, the Butterworth turns out to be the simpler design, even though it has a higher filter order, since a delay equalizer is not needed.

Let's choose a design with $R_s = 1$, $L_1 = 0.445$, $C_2 = 1.247$, $L_3 = 1.8019$, $C_4 = 2.0$, $L_5 = 1.8019$, $C_6 = 1.247$, $L_7 = 0.445$ and $R_T = 1$. The unnormalized values are shown in the circuit of **Figure 14-32a**. We note that we meet the gain specification as well as the group delay specification (**Figure 14-32b**) using this filter, without any additional group delay equalization.

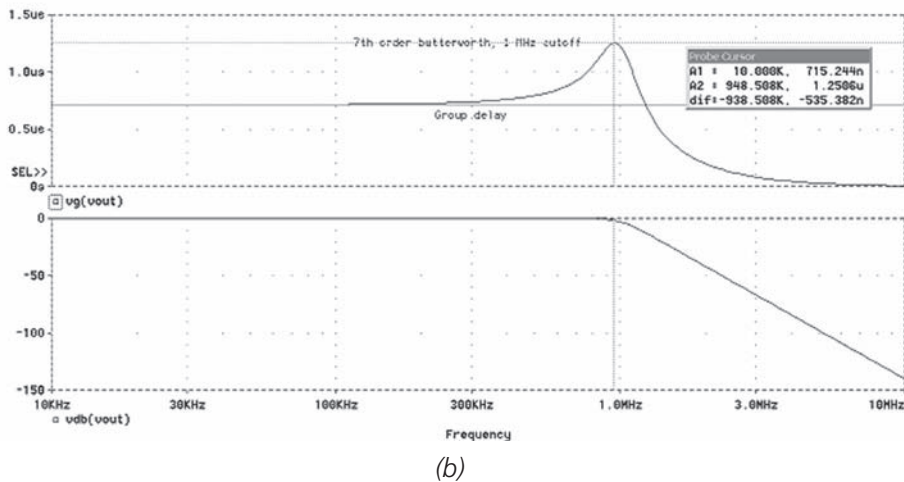
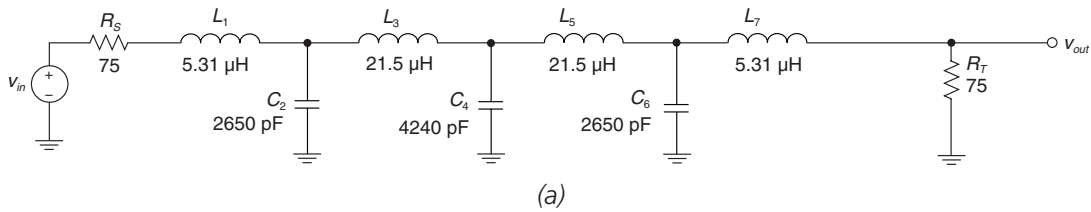


Figure 14-32: $N = 7$ Butterworth ladder filter response. (a) Circuit. (b) Peak-peak group delay variation is less than 1000 ns (approximately 535 ns). Note that a gain of +2 to compensate for the resistive divider is assumed following the filter.

Chapter 14 Problems

Problem 14.1

Design a seventh-order Butterworth low-pass filter with cutoff frequency 4.5 MHz. Implement this using an LC ladder with a 50Ω source impedance. Assume that you follow the filter with a gain of +2 to compensate for the -6dB loss due to the source and termination resistances.

Problem 14.2

Calculate the attenuation of the 4.5-MHz filter in Problem 14.2 at 13.5 MHz.

Problem 14.3

Design a fifth-order Butterworth filter with cutoff frequency 50 kHz, using Sallen-Key sections.

Problem 14.4

Compare the step responses of a seventh-order 4.5-MHz Butterworth filter and a seventh-order Chebyshev filter with cutoff frequency 4.5 MHz. Comment on the risetime and overshoot of each filter.

Problem 14.5

Design a sixth-order Chebyshev filter with 0.5dB ripple and cutoff frequency of 3.58 MHz. Find the peak-peak group delay variation in this filter.

Problem 14.6

An $N = 5$ Butterworth filter design with equal source and termination resistances is shown below. The -3dB frequency of the filter shown is 1 radian/second (0.16 Hz).

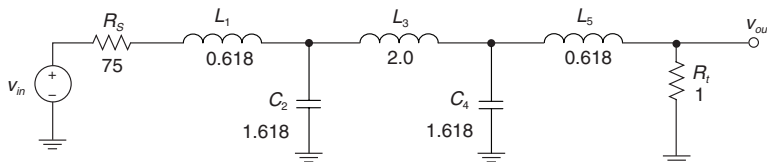


Figure 14-33: $N = 5$, 1 radian/second Butterworth low-pass filter.

- (a) Calculate the frequency (in radians/second) at which the output attenuation drops to -40dB for the filter as built above.

- (b) Now, assume a source and termination resistance of 75Ω . You now need a filter with a -3dB cutoff at 10 MHz (not radians/sec!). Find the unnormalized Butterworth values and sketch the circuit.
- (c) Sketch the group delay vs. frequency of the 10-MHz filter. In your sketch, include at least the following data points: DC, 5 MHz, 10 MHz and 15 MHz. Indicate the frequency at which the group delay peaks.

Problem 14.7

A low-pass filter prototype with -3dB cutoff frequency has a group delay at DC of 1 second. The prototype filter is then transformed and a filter with the same type and order is created, but now with a -3dB cutoff frequency of 4.5 MHz. What is the group delay at DC of the new filter?

References

- Balch, Brent, "A Simple Technique Boosts Performance of Active Filters," *EDN*, November 10, 1988, pp. 277–286.
- Blinchikoff, H., and Zverev, A., *Filtering in the Time and Frequency Domains*, John Wiley, 1976.
- Burton, L. T., and Treleaven, D., "Active Filter Design Using Generalized Impedance Converters," *EDN*, February 5, 1973, pp. 68–75.
- Chambers, William, "Know Your Options and Requirements when Designing Filters," *EDN*, August 5, 1991, pp. 129–138.
- Corral, C., "Designing elliptic filters with maximum selectivity," *EDN*, May 25, 2000, pp. 101–109.
- Corrington, Murlan S., "Transient Response of Filters," *RCA Review*, September 1949, vol. 10, no. 3, pp. 397–429.
- Downs, Rick, "Vintage Filter Scheme Yields Low Distortion in New Audio Designs," *EDN*, November 7, 1991, pp. 267–272.
- Steer, Robert, Jr., "Antialiasing Filters Reduce Errors in A/D Converters," *EDN*, March 30, 1989, pp. 171–186.
- Tow, J., "A Step-by-Step Active-Filter Design," *IEEE Spectrum*, December 1969, pp. 64–68.
- Williams, A., and Taylor, F., *Electronic Filter Design Handbook*, McGraw-Hill, 1988.
- Yager, Charles, and Laber, Carlos, "Create a High-Frequency Complex Filter," *Electronic Design*, April 13, 1989, pp. 123–127.
- Zverev, Anatol, *Handbook of Filter Synthesis*, John Wiley, 1967.

Review of Passive Components and a Case Study in PC Board Layout

In This Chapter

- Some of the subtleties of passive components including construction techniques of these devices and parasitic effects are considered. We'll cover some details about resistors, capacitors and inductors. Then, we'll use what we've learned in an illustrative discussion of PC board layout issues.

Resistors

At first blush, a resistor is a resistor is a resistor. However, we'll now delve into some of the subtleties of these devices. The impedance of an ideal resistor is not dependent on operating frequency,¹ and is:

$$Z_{\text{resistor,ideal}} = R \quad [15-1]$$

A real-world resistor (**Figure 15-1**) includes a parasitic inductance due to the geometry of the lead length, and a parasitic capacitance across the resistor. The impedance of the resistor, including these parasitic elements, is:²

$$Z_{\text{resistor,real}}(s) = \frac{Ls + R}{LCs^2 + RCs + 1} = \frac{R \left(1 + \frac{L}{R}s \right)}{LCs^2 + RCs + 1} \quad [15-2]$$

We can put this in “ $j\omega$ ” form by making the substitution $s = j\omega$ resulting in:

$$Z_{\text{resistor,real}}(j\omega) = \frac{j\omega L + R}{(1 - \omega^2 LC) + j\omega RC} \quad [15-3]$$

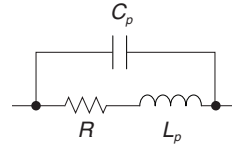
¹ Another way to look at this is that the current in an ideal resistor and a voltage across this ideal resistor are in phase.

² One way to sanity-check this result is to consider the limit as $L \rightarrow 0$ and $C \rightarrow 0$; we want the impedance to be exactly R and this is indeed the case.

For a large-valued resistor,³ the RC time constant dominates. This is because a large resistor will swamp out the value of the parasitic inductance. For a low-valued resistor the L/R time constant dominates because the resistor effectively shorts out the parasitic capacitance. The magnitude of the impedance of the real-world resistor is:

$$|Z_{\text{resistor,real}}| = \sqrt{\frac{(\omega L)^2 + R^2}{(1 - \omega^2 LC)^2 + (\omega RC)^2}} \quad [15-4]$$

Figure 15-1: Resistor R showing parasitic elements series inductance L_p and parallel capacitance C_p .



In **Figure 15-2**, we see the impedance of a resistor with $R = 1 \text{ M}\Omega$, $C = 0.2 \text{ pF}$ and $L = 10 \text{ nH}$. Note that for this relatively large resistor ($R \gg Z_o = 223\Omega$), parasitic capacitive effects dominate; the impedance rolls off at frequencies above approximately 1 MHz.

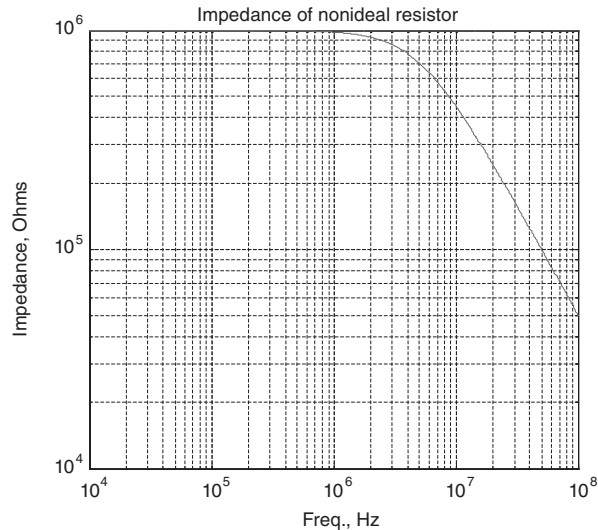


Figure 15-2: Impedance of nonideal resistor with $R = 1 \text{ M}\Omega$, $C = 0.2 \text{ pF}$ and $L = 10 \text{ nH}$.

³ We can see that “large-valued” in this case is large enough so that $RC \gg L/R$, or equivalently $R \gg \sqrt{\frac{L}{C}}$. The term $\sqrt{\frac{L}{C}}$ comes up over and over again in RLC circuits and transmission lines and is called the *characteristic impedance* Z_o of this circuit.

In **Figure 15-3**, we see the impedance of a low-valued resistor with $R = 10\Omega$, $C = 0.2$ pF and $L = 10$ nH. Note that parasitic inductive effects dominate; the impedance increases at frequencies above approximately 10 Mrad/sec.

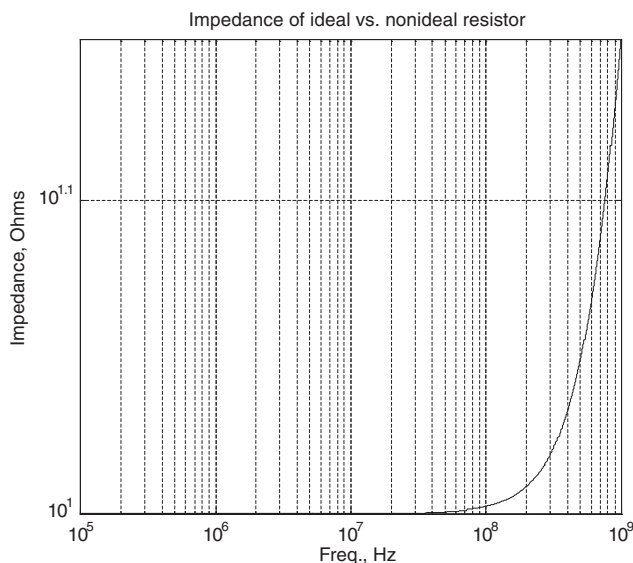


Figure 15-3: Impedance of nonideal resistor $R = 10\Omega$, $C = 0.2$ pF and $L = 10$ nH.

It's difficult to quantify exactly the values of parasitic elements, but in standard through-hole resistors you might expect fractions of a picofarad of parasitic capacitance, and a few nanohenries of parasitic inductance.⁴ You need to consider such parasitics as series inductance and parallel resistance in high-frequency circuits.

Comments on Surface-Mount Resistors

You can miniaturize your circuit and reduce somewhat the effects of parasitic inductance by using surface-mount resistors. Surface-mount resistors come in a variety of sizes, ranging from "0201" size, through "0402" size and all the way up to "2512" size. The resistor size numbers indicate the length and width of the resistor. For instance, an 0805 resistor is 0.08" in length and 0.05" in width.

There is a trade-off to be made with regard to resistor size and wattage and working voltage rating, as shown in **Table 15-1**. For instance, a longer resistor will have a higher working voltage rating, due to a higher voltage breakdown.

⁴ A very rough rule of thumb for the inductance of component leads above a ground plane is 10 nanohenries per centimeter of lead length. So, it behooves you to keep lead lengths short if you want to minimize parasitic inductance. Of course, you can test your resistor using an impedance analyzer (such as the Hewlett Packard HP4192) and extract the parameters for your device.

Table 15-1: Comparison of surface mount resistor ratings.

Resistor	Size	Typical wattage rating ⁵	Typical working voltage rating ⁶
0201	0.02" × 0.01"	50 mW	15V
0402	0.04" × 0.02"	50 mW – 62.5 mW	50V
0603	0.06" × 0.03"	62.5 mW – 100 mW	75V
0805	0.08" × 0.05"	100 mW – 250 mW	100V
1206	0.12" × 0.06"	125 mW – 250 mW	200V
1210	0.12" × 0.10"	250 mW – 333 mW	200V
1812	0.18" × 0.12"	500 mW	200V
2010	0.20" × 0.10"	500 mW	200V
2512	0.25" × 0.12"	1000 mW	250V

Comments on Resistor Types

As the designer, you also have decisions to make regarding the type of resistor you put in your circuit. For instance, do you choose carbon composition, carbon film, metal film, wire-wound, or some other type of resistor?

Carbon composition resistors, sometimes called *carbon comp* are old-style resistors that have been used for years and years in electronics. The main advantage of carbon composition resistors is their ability to withstand high current transient surges. They do have the disadvantage of a high temperature coefficient of resistivity. Remember that the resistance of a resistor varies with temperature, and that the resistance can be expressed as:

$$R(T) = R_o (1 + \alpha(T - T_o)) \quad [15-5]$$

- $R(T)$ = resistance at your operating temperature
- R_o is the reference resistance at temperature T_o
- α is the temperature coefficient of resistivity

Carbon composition resistors also have a tendency to drift in value with time, especially if they are overstressed with high currents. Carbon comp resistors have largely been replaced in modern electronics by metal film and carbon film resistors. Film resistors have the advantage of a lower temperature coefficient of resistivity. They are, however, somewhat more susceptible to damage by electrical overloads.

Wirewound resistors are largely used where high-wattage capability is needed. They do suffer, however, from a large series inductance due to the way they are manufactured with wound wires.

A comparison of resistor types is given in **Table 15-2**.

⁵ There is some variation in wattage rating from manufacturer to manufacturer, so be sure to check the particular device datasheet.

⁶ Again, use these numbers for comparison purposes only, and check the specific manufacturer.

Table 15-2: Comparison of resistor types.

Resistor	Typical power rating	Temperature coefficient	Comments
Carbon composition	0.25W ~ 2W	> 1000 ppm/°C	Old-style resistors. Typically replaced by carbon film or metal film in new designs. Poor long-term stability and temperature coefficient.
Carbon film		Typically -50 ppm/°C to -1000 ppm/°C	
Metal film		Typically +50 ppm/°C to +300 ppm/°C	Low noise
Wirewound	Typically > 5W	Typically +100 ppm/°C	Typically used for high wattage resistors. Be careful of high parasitic inductance

Capacitors

Just as resistors suffer from parasitic components, so do capacitors. A model of a real-world capacitor constructed as a parallel plate filled with a dielectric is shown⁷ in **Figure 15-4a**. The resistance R_s is the series resistance of the leads. The parallel plate is filled with a dielectric that has a finite electrical conductivity.⁸ This results in a dielectric resistance R_d that is in parallel with the desired capacitance C , as shown in **Figure 15-4b**.

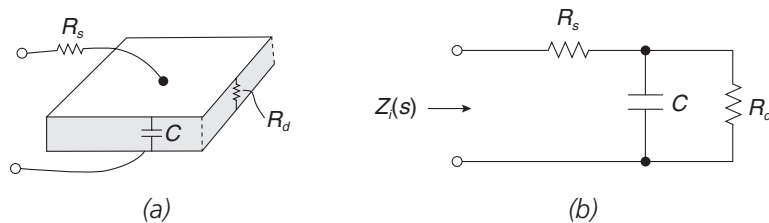


Figure 15-4: Capacitor showing parasitic elements. (a) Parallel-plate capacitor filled with dielectric with finite electrical conductivity. This capacitor has series resistance R_s and dielectric resistance R_d that is in parallel with the lumped capacitance C . (b) Electrical model.

The input impedance to the real-world capacitor is:

$$Z_i(s) = R_s + \frac{R_d}{R_d C s + 1} \quad [15-6]$$

⁷ In this example, we've ignored the series inductance for simplicity.

⁸ For further information on the lossy capacitor see, e.g., Markus Zahn, *Electromagnetic Field Theory: A Problem Solving Approach*, Krieger reprint 1987, pp. 184–194. Note that in this initial model we don't include the effects of series inductance. We'll consider this in more detail later.

We can expand this result to find the real and imaginary parts of the input impedance as follows:

$$\begin{aligned}
 Z_i(j\omega) &= R_s + \frac{R_d}{j\omega R_d C + 1} \\
 &= R_s + \frac{R_d(1 - j\omega R_d C)}{1 + \omega^2 R_d^2 C^2} \\
 &= \left[R_s + \frac{R_d}{1 + \omega^2 R_d^2 C^2} \right] - j \left[\frac{\omega R_d^2 C}{1 + \omega^2 R_d^2 C^2} \right]
 \end{aligned}
 \tag{15-7}$$

The first term (the real part) is sometimes called the *equivalent series resistance* of the capacitor, or:

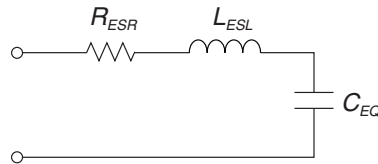
$$R_{ESR} = \left[R_s + \frac{R_d}{1 + \omega^2 R_d^2 C^2} \right] \tag{15-8}$$

Note that the ESR decreases as frequency increases. The equivalent capacitance is:

$$C_{eq} = C \left[1 + \frac{1}{\omega^2 R_d^2 C^2} \right] \tag{15-9}$$

A simplified model of the capacitor showing the equivalent series resistance is shown in **Figure 15-5**. To this model we've also added an equivalent series inductance (L_{ESL}). The value of the series inductance depends on the geometry of the internal construction of the capacitor, as well as the lead length of the device as it is connected in the circuit. A ballpark rule-of-thumb for ESL of a capacitance is 10 nanohenries per centimeter of lead length. The ESL of some electrolytic capacitors can be somewhat higher if the capacitor is fabricated with wound foil inside a can.

Figure 15-5: Capacitor with equivalent series resistance (R_{ESR}) and equivalent series inductance (L_{ESL}).



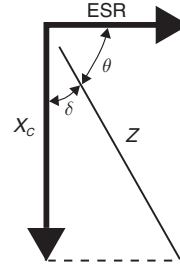
The dissipation factor (DF) is another figure-of-merit often found on capacitor datasheets. The dissipation factor is given by:

$$DF = \frac{R_{ESR}}{\omega C} \tag{15-10}$$

We note that the dissipation factor is the inverse of the Q of the capacitor. If we plot ESR, capacitive reactance (X_C), and total capacitor impedance (Z) as in **Figure 15-6**, we see that there is a phase angle between the capacitive reactance and the impedance of the capacitor. The dissipation factor is the tangent of this angle, or:

$$DF = \tan(\delta) \tag{15-11}$$

Figure 15-6: Plot illustrating the dissipation factor. The dissipation factor is the tangent of the angle between the impedance of the capacitor X_c and the overall impedance Z .



The impedance of an ideal capacitor is:

$$Z_{cap,ideal} = \frac{1}{j\omega C} \quad [15-12]$$

For a real-world capacitor (ignoring dielectric loss) the impedance is:

$$Z_{cap,real} = \frac{1}{j\omega C} + R + j\omega L = \frac{(1 - \omega^2 LC) + j\omega RC}{j\omega C} \quad [15-13]$$

The magnitude of the impedance is:

$$|Z_{cap,real}| = \frac{\sqrt{(1 - \omega^2 LC)^2 + (\omega RC)^2}}{\omega C} \quad [15-14]$$

A impedance plot of an electrolytic capacitor, comparing the ideal with the actual impedance, is shown in **Figure 15-7** for $C = 100 \mu\text{F}$, $L = 25 \text{ nH}$ and $R = 0.01\Omega$. We note that at frequencies above 1 MHz the impedance of the capacitor looks inductive.

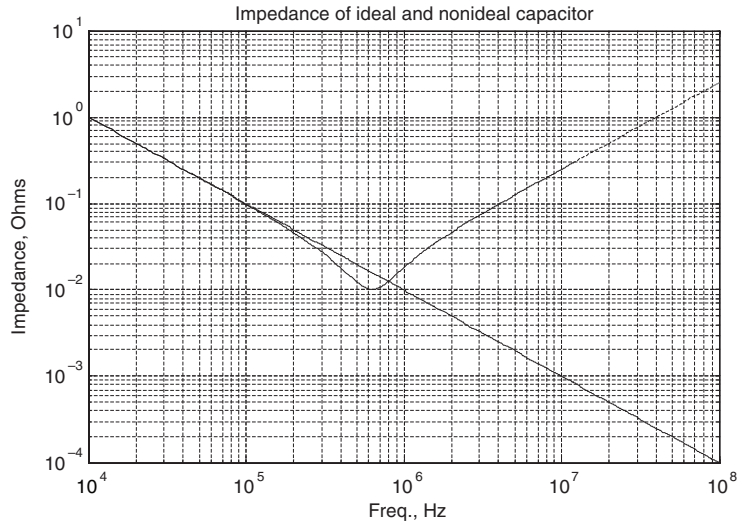


Figure 15-7: Impedance plot of electrolytic capacitor with $C = 100 \mu\text{F}$, $L = 25 \text{ nH}$ and $R = 0.01\Omega$. Dotted line is the impedance of an ideal 100- μF capacitor.

Inductors

The impedance of an ideal inductor is:

$$Z_{\text{inductor,ideal}} = j\omega L \quad [15-15]$$

For a real-world inductor the impedance is modified by the resistance of the copper wire⁹ and the inter-winding capacitance (**Figure 15-8**). This impedance is:

$$Z_{\text{cap,real}} = \frac{j\omega L + R}{(1 - \omega^2 LC) + j\omega RC} \quad [15-16]$$

The magnitude of this impedance is:

$$|Z_{\text{inductor,real}}| = \sqrt{\frac{(\omega L)^2 + R^2}{(1 - \omega^2 LC)^2 + (\omega RC)^2}} \quad [15-17]$$

A impedance plot of an electrolytic capacitor, comparing the ideal with the actual impedance, is shown in **Figure 15-9** for $L = 100 \mu\text{H}$, $C = 25 \text{ pF}$ and $R = 0.1 \Omega$. Note that the self-resonant frequency of the inductor (at approximately 20 Mrad/sec.) is clearly shown. Above the self-resonant frequency, the impedance of this inductor is capacitive.

Figure 15-8: Inductor showing parasitic elements of series resistance R and inter-winding capacitance C .

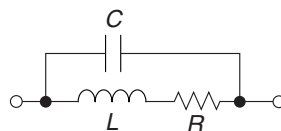
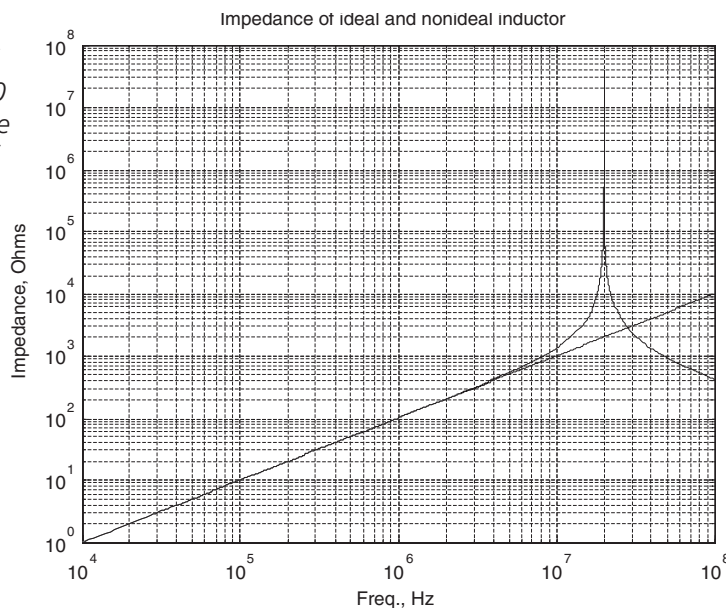


Figure 15-9: Impedance plot of real-world inductor with $L = 100 \mu\text{H}$, $C = 25 \text{ pF}$ and $R = 0.1 \Omega$. The dotted line is the impedance of an ideal $100\text{-}\mu\text{H}$ inductor.



⁹ This analysis ignores core losses (if any), and other high-frequency effects such as current crowding due to skin effect.

Discussion of Printed-Circuit Board Layout Issues

Printed circuit (PC) board layout and routing is a task that is sometimes left to the last minute in a design cycle. Doing a good PC board layout requires attention to many details, including:

- Knowledge of where you want the high frequency, high current, or sensitive circuitry to be.
- Some information on component limitations.
- Information on noise sources.
- Real-world constraints, such as PC board form-factor and location of connectors and mounting holes.
- Other constraints such as PC board design rules mandating minimum trace widths, trace-to-trace spacing, and the like.

Following is a discussion of some of these design issues.

Power supply bypassing

The need for power supply bypassing from integrated circuits arises from the fact that there is no such thing as a perfect, zero-impedance ground. Consider the model of an integrated circuit IC_1 in **Figure 15-10**, the details of which are unimportant for purposes of this discussion. The IC draws DC power from the supply (V_{supply}) through wires or ground and power planes. The series inductance and resistance of the interconnection to the supply is shown.

The integrated circuit IC_1 draws a fast switching current with a high di/dt (modeled as current source $i(t)$). The hope in bypassing the IC is that proper selection and placement of the bypass capacitor C_B will force transient currents to circulate locally near the integrated circuit and hence voltage transients on the supply lines will be limited. Of course, we want the DC component of the IC_1 current to travel back to the power supply. However, if we send fast current pulses back to the power supply, we will induce voltages on the power supply lines to IC_2 and IC_3 due to distributed resistance and inductance.

The key to selection and placement of bypass capacitor C_B is to choose a capacitor that is sufficiently sized to do the job, and to place it in close proximity to the power and ground pins of IC_1 . This will minimize the inductance of the bypass path and ensure that the high-frequency switching currents circulate locally near the IC .

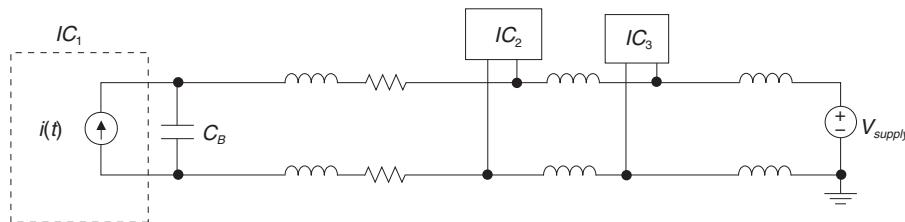


Figure 15-10: Model showing integrated circuit drawing current, and its bypass path. C_B is the bypass capacitor for IC_1 .

One way to reduce the impedance of current return paths is to use a ground plane, discussed in the following section.

Ground planes

A ground is a return path for current. It is desirable that this return path should have as low an impedance as possible, to reduce transient-induced voltage drops and electromagnetic emissions. In the world of two-layer PC boards, it's difficult to have a dedicated ground plane since you generally want a couple of PC board layers available for routing signals. In multilayer boards, it's easy to dedicate an unbroken ground plane on an internal PC board layer as ground.

The use of a ground plane helps to reduce the inductance of signal-carrying traces on the PC board. One technique is to have high current and high di/dt traces directly above an unbroken ground plane. You can also make the traces wide if you want to reduce the inductance.

PCB trace widths

PCB traces must be sized appropriately (both in width and thickness, or copper weight¹⁰) to carry the current that you need without excessive temperature rise. A rule of thumb is that a 10-mil-wide, 1-ounce PC board trace can carry in excess of 500 mA with a 20°C temperature rise above ambient. An estimate of the current-carrying capability for 20°C temperature rise of PC board traces is shown in **Figure 15-11**. The fusing current (**Figure 15-12**) for PC traces is higher.

Table 15-3: PC board copper weight vs. thickness.

Copper weight	Copper thickness (inches/mils)
½ oz.	0.0007" (0.7 mils)
1.0 oz.	0.0014" (1.4 mils)
2.0 oz.	0.0028" (2.8 mils)

¹⁰ “Copper weight” tells you how thick the PC board trace is. Typical low-power analog boards use ½-ounce or 1-ounce copper. High-power boards may use 2-ounce copper or higher.

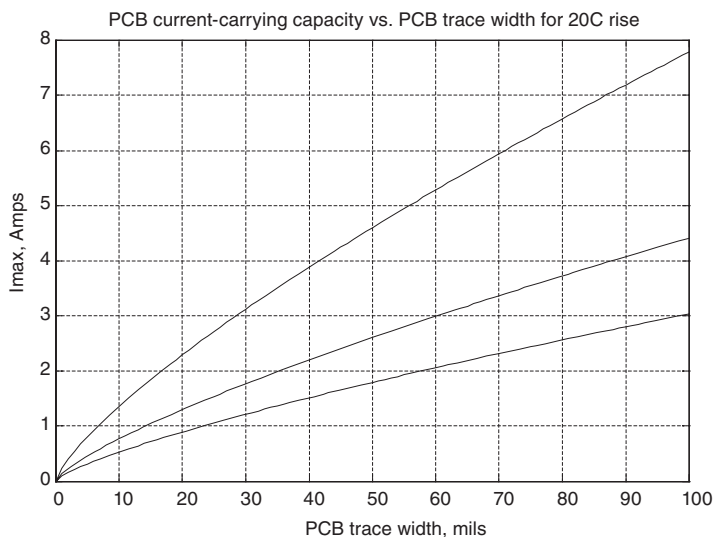


Figure 15-11: Approximate current-carrying capability of 0.5 oz., 1.0 oz. and 2.0 oz. PCB traces with 20°C temperature rise.¹¹ The upper trace is 2 oz. copper; the lower trace is 0.5 oz. (Note: 1 mil = 0.001")

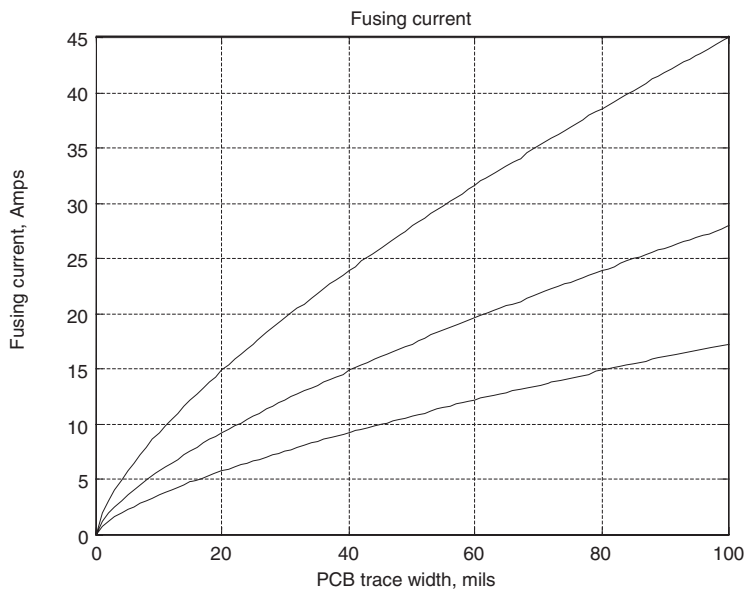


Figure 15-12: Approximate fusing current of 0.5 oz., 1.0 oz. and 2.0 oz. PCB¹² traces.

¹¹ From Douglas Brooks, reference at the end of this chapter.

¹² From Douglas Brooks, "Fusing Current" reference at the end of this chapter

Approximate Inductance of a PC Board Trace Above a Ground Plane

The inductance of a PC board trace above a ground plane can be roughly calculated by assuming a microstrip configuration. For a microstrip line of length l , width w and strip to strip spacing $d \ll w$ (**Figure 15-13a**) the inductance is (very roughly¹³):

$$L \approx \mu_o \frac{lh}{w} \quad [15-18]$$

Using this approximation for a line with $w = 0.01''$ (0.0254 cm) and $h = 0.005''$ (0.0127 cm) we estimate an inductance of 6.3 nanohenries per centimeter of length. A 2D finite element analysis¹⁴ (**Figure 15-13b**) estimates this inductance to be somewhat lower at approximately 3.9 nanohenries per centimeter of length.

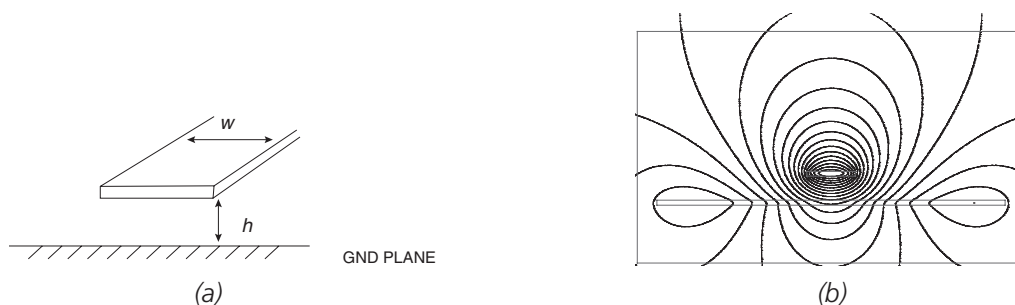


Figure 15-13: 2D FEA model of $w = 0.01''$ PC board trace $0.005''$ above a ground plane. (a) Geometry. (b) 2D FEA model.

Example 15.1: Design case study—high-speed semiconductor laser diode driver

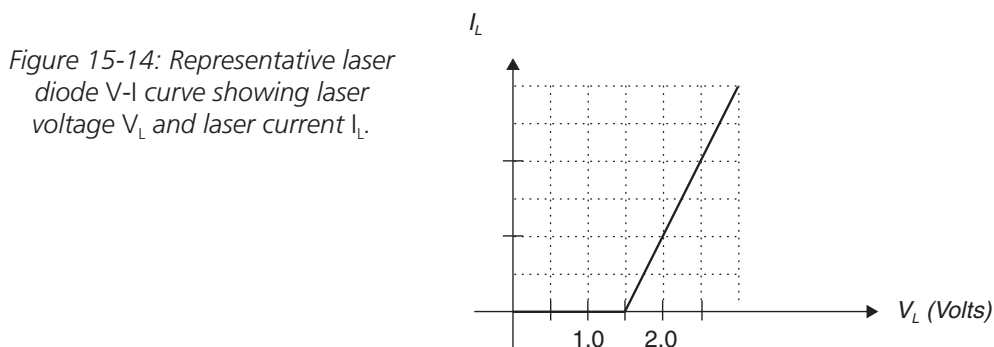
This section considers the design, analysis, and PC board layout of a high-speed switching semiconductor laser diode system, which may be used as a modulated infrared (IR) light source. Direct modulation is a method by which the laser light power output of a semiconductor laser diode is changed by varying the diode current. To use a diode as a high-speed modulated light source, the laser is biased with a small DC current near the *lasing threshold* and a modulation current is superimposed. The light power output of the semiconductor diode is proportional to the laser current in excess of the threshold current. The direct modulation

¹³ Note that this approximation becomes less and less accurate as the trace height h increases above the ground plane. So, use this for ballpark estimates only. For more detailed calculations for inductances of all kinds of geometries, see Frederick Grover's excellent reference *Inductance Calculations*, reference given at the end of this chapter. The author gratefully thanks Prof. Dave Perreault from MIT for recommending this book when we were both grad students.

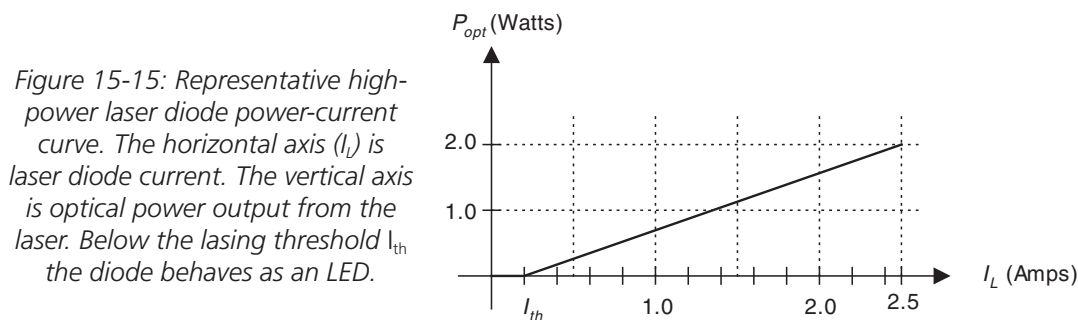
¹⁴ Plots and analysis done with Finite Element Method Magnetics (FEMM), a finite element package created by Dr. David Meeker at Foster-Miller.

method is used for laser communication, for fiber-optic links, for industrial applications such as material cutting, and in such commercial products as compact-disk players and medical laser printers.

Electrically, a semiconductor laser behaves like a diode, with a V/I curve shown in **Figure 15-14**. Since the semiconductor diode is made from gallium arsenide (GaAs) rather than silicon, the voltage “knee” when the diode turns on is approximately 1.5V.



Under normal operation, the diode is driven by a current source so that the diode current remains constant even if the diode voltage drifts with time and temperature. The optical power output vs. diode current is shown in **Figure 15-15** for a high-power laser diode. For very low currents, the diode does not lase and there is very little optical power. (In fact, for current $< I_{th}$, the laser behaves like an LED and there is some very small amount of optical power emitted.) Once the diode current is increased to a value known as the threshold current (I_{th}) the diode begins lasing, and the optical power output is proportional to the current in excess of the laser threshold current. For a 2W laser diode, the operating laser current is approximately 2.5A as shown in **Figure 15-15**. If the laser current is increased further, the laser may be damaged by a process known as catastrophic optical damage (COD) where excess heating destroys¹⁵ the laser-emitting area.



¹⁵ This curve is representative of one particular high-power laser diode used by the author; there are other lasers with different power levels and operating currents.

Semiconductor laser diodes are inherently fast devices. The intrinsic lasing processes may be modulated at very high rates by variation of the injected current. For representative diodes, the laser power transfer function (optical power output due to current excitation) is flat out to several hundreds of megahertz, or even higher (**Figure 15-16**), depending on the details of the diode construction and the current bias level. The resonance near 10^{10} radians/second is due to quantum relaxation processes. Therefore, the high light modulation speed indicated may be achieved in practice if the laser current is changed sufficiently fast. This next leads us to consider how to switch laser current with high current and fast risetimes.

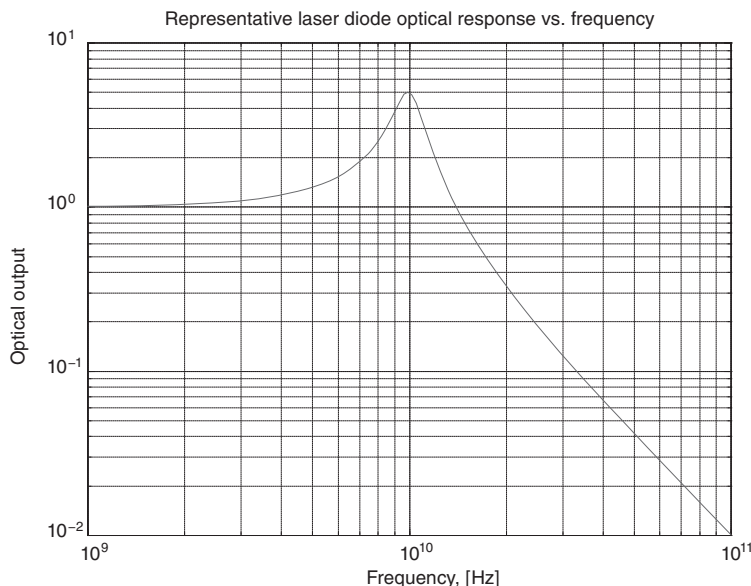


Figure 15-16: Representative laser diode intrinsic light output frequency response with laser resonance at 1 GHz. Horizontal axis—frequency, Hz.

Driver implementation

One possible circuit topology suitable for driving a laser diode is shown in **Figure 15-17**. The laser is fed by two DC current sources I_{BIAS} and I_{th} , corresponding to a laser PEAK current and THRESHOLD currents. When V_{B1} is LOW and V_{B2} is HIGH, Q_1 is OFF and Q_2 is ON, the total current in the laser diode is $I_{PK} + I_{th}$. The resistor in the collector of Q_1 dissipates power so Q_1 will not be damaged.

The author was responsible for the design of a semiconductor diode laser modulator capable of delivering 2.5A pulses to a low impedance load with risetime and falltime of less than 20 ns. The purpose of the circuit board was to drive semiconductor diode lasers for high-speed printing.¹⁶

¹⁶ For more details on the design, see Marc Thompson and Martin Schlecht, "High Power Laser Diode Driver Based on Power Converter Technology," *IEEE Transactions on Power Electronics*, vol. 12, no. 1, January 1997, pp. 46–52 and U.S. Patent 5,444,728 (issued 8/22/95).

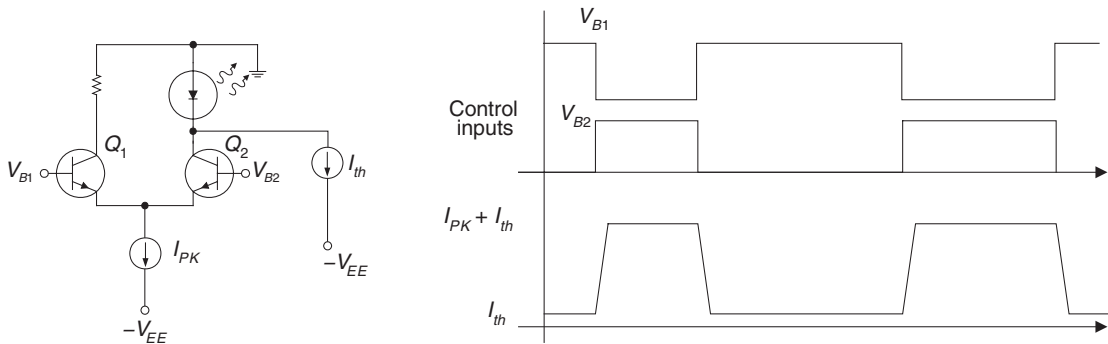


Figure 15-17: Laser driver circuit implemented as an emitter-coupled switch, and waveforms. The top traces are the control inputs V_{B1} and V_{B2} . The bottom trace is the laser current.

There were several design challenges inherent in this design. First, the laser signal is a high-current, fast-risetime set of current pulses with any repetition rate from DC up to 10 MHz, with any duty cycle. This means that extreme care must be taken to ensure low-inductance path from the switching elements on the PC board to the laser. The design of the switch on the PC board presents interesting thermal problems as well.

A simplified schematic of the switching transistor array is shown in **Figure 15-18**. In order to provide low-inductance paths as well as good thermal management, the fast switch was broken up into a dozen smaller emitter-coupled switches, each pair implemented with a pair of 2N2222 transistors. Note that in emitter coupled pair Q_{1a} and Q_{1b} only one transistor is on at a time; when DRIVE is HIGH and $\overline{\text{DRIVE}}$ is LOW, Q_{1a} is ON and Q_{1b} is OFF (and hence the laser is off, and idling at the threshold current I_{th}). When DRIVE is LOW and $\overline{\text{DRIVE}}$ is HIGH, Q_{1b} is ON and Q_{1a} is OFF and the total laser current is $I_{th} + I_{PK}$.

The critical high-speed and high-current switching paths are highlighted¹⁷ in bold. Each of the transistor arrays switches up to a maximum of more than 200 mA. It's mandatory to keep the interconnection inductance between transistors and to the laser diode low in order that the transistor arrays can switch as fast as they are capable. Remember from previous chapters that emitter-coupled switches are inherently fast, provided that you provide sufficient base drive capability.¹⁸

Resistors R_{B1} , R_{B2} up to R_{B12} up to are low-valued *ballast* resistors and ensure that the transistor pairs share the current among them equally.¹⁹

¹⁷ When doing a PC board layout, the traces in bold are good candidates to be implemented as wide traces over an unbroken ground plane, to reduce parasitic inductance.

¹⁸ See, e.g., Chapter 10 where we showed that the switching speed of the emitter-coupled pair for signal transistors to be a few nanoseconds. Of course, this assumes that we have a good PC board layout so that parasitic inductances don't slow things down significantly.

¹⁹ The transistors in this design aren't matched, and we want each transistor pair to shoulder an equal value of the load. The ballast resistors, on the order of 1Ω , forces sharing between transistor pairs.

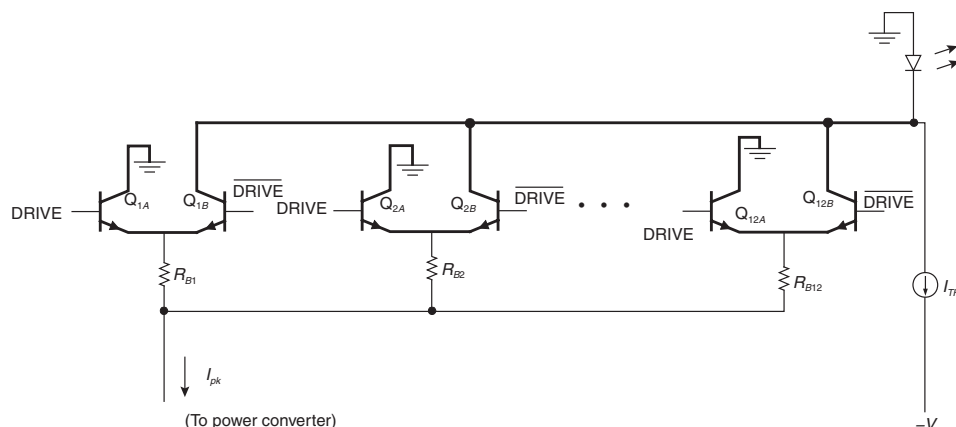
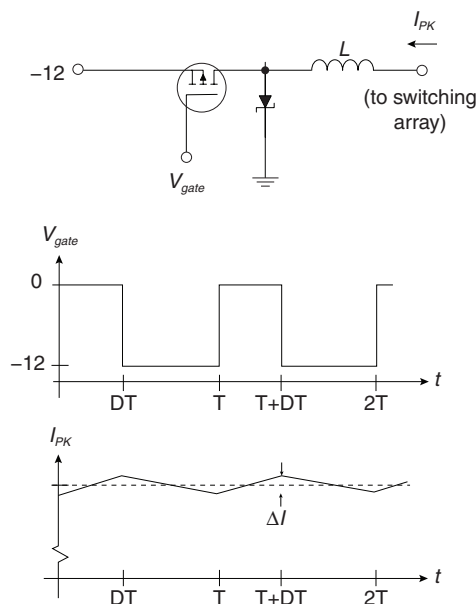


Figure 15-18: Switching transistor array capable of fast risetime switching of up to 2.5A. The array is comprised of 12 pairs of high-speed switching transistors. The critical high-speed and high-current switching paths are shown in bold.

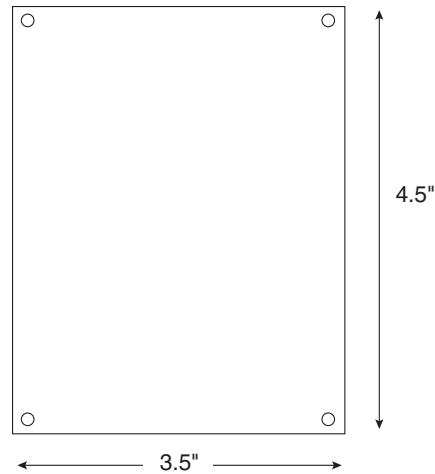
Another design challenge was that the current rating of the -12V power supply was only 1A, and we want to deliver 2.5A to the lasers. Therefore, a DC/DC converter was needed to step down the voltage and step up the current. A simplified implementation of this is shown **Figure 15-19**. A detailed discussion of the DC/DC converter is beyond the scope of this book, but this circuit steps DOWN the voltage and steps UP the current. Hence we draw less current from the -12V power supplies than delivered to the lasers. We note that this is a high-speed switching circuit and hence we need to take special care in the layout of the MOSFET and diode. The power MOSFET is switched on and off at a high frequency with a variable duty cycle to regulate the current, as shown.

Figure 15-19: Simplified schematic of the DC/DC converter.



In this design, the form factor of the PC board was mandated to be $3.5" \times 4.5"$, as this design was replacing a pre-existing design and the PC boards had to be backward-compatible. The form factor and mounting holes are shown in **Figure 15-20**.

Figure 15-20: Form factor of PC board, $3.5" \times 4.5"$, showing mounting holes.



The connector locations (**Figure 15-21**) were also set prior to the layout. The connectors are as follows (clockwise from bottom left):

- **Power:** +12V @ 200 mA; -12V @ 1A, and 2 ground pins.
- **Laser diode connection:** A microstrip cable was soldered directly to the PC board to provide a low-impedance path to the lasers.
- **Trigger signal:** This is a TTL-level signal that turns the laser diode ON and OFF. When the trigger signal is high, the laser is ON. Repetition rates for the TTL signal is from DC up to 10 MHz.
- **Monitor:** This connector is used to buffer and amplify a photodiode signal used to monitor the optical power output of the laser.
- **Shutdown:** Another TTL-level signal which is used to completely shut down the laser.

Next, the real estate for the various PC board traces was allocated as in **Figure 15-22**. We note that the high-current and high-speed circuitry is segregated from the low-level analog instrumentation circuitry. Furthermore, the PC board was multilayer, ensuring that an unbroken ground plane could be used under the high-speed circuitry. A breakdown of the PC board layers is as follows:

- Top layer: Analog signals
- Internal layer #1: GND
- Internal layer #2: -12V
- Bottom layer: Analog signals, +12V

Figure 15-21: PC board connector locations.

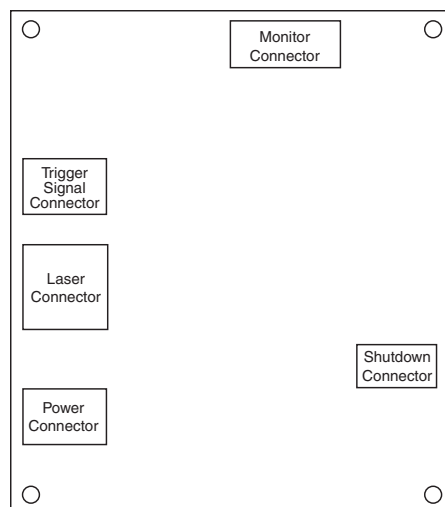
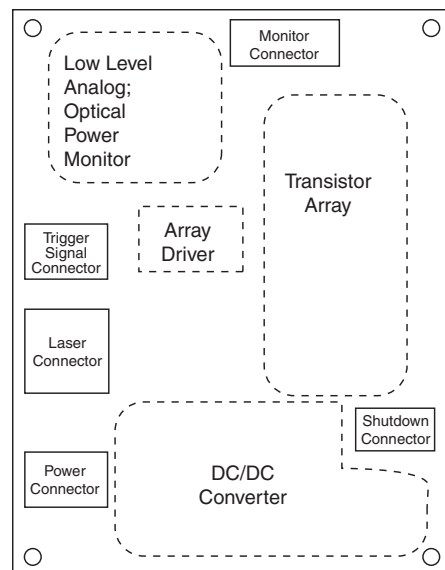


Figure 15-22: PC board allocation of real estate.



An internal layer was dedicated to -12V since there were significant switching currents drawn from the DC/DC converter.

A photograph of the resultant PC board is shown in **Figure 15-23**. An oscilloscope photograph of the laser light output²⁰ is shown in **Figure 15-24**. We note that the laser is switching 2W peak-peak, corresponding to a switched current of 2.5A, peak-peak. The risetime and falltime is less than 20 ns.

²⁰ The light output was measured using an extremely fast photodetector. Since the laser is an inherently fast device, the light output shape is representative of the shape of the current pulses to the laser.

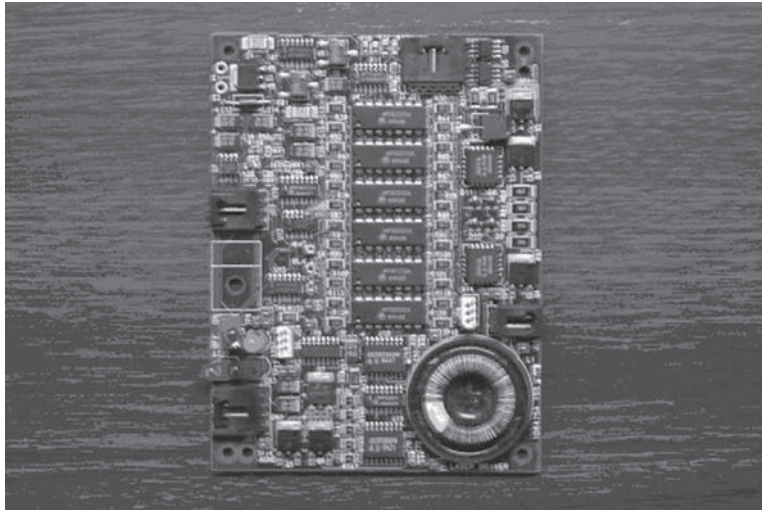


Figure 15-23: PC board showing top side (component side) final layout. Connection to laser diode is not shown.

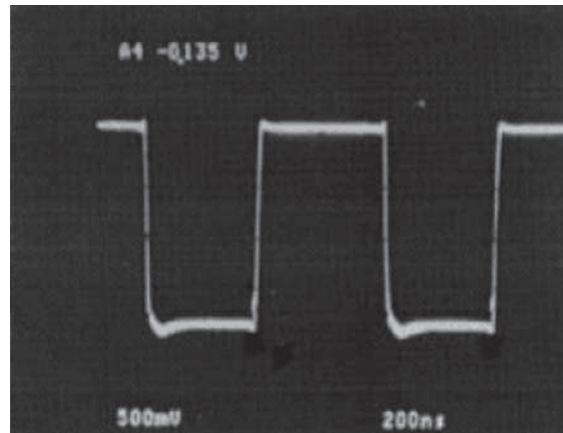


Figure 15-24: Scope photograph showing risetime and falltime of laser optical power. Horizontal: 200 ns per division. Vertical: 2W full scale. The resultant laser light 10–90% risetime and falltime is less than 20 ns.

Chapter 15 Problems

Problem 15.1

This problem concerns real-world capacitors. Every capacitor has some parasitic elements; that is, there is no such thing as an “ideal capacitor.” For instance, the leads that connect a capacitor to your circuit have a parasitic inductance that is in series with the capacitor. This parasitic series inductance affects how the capacitor operates at high frequency. Also, the capacitor has some internal resistance, also in series with the capacitance, known as the equivalent series resistance (ESR).

- (a) Draw the lumped circuit model for the real-world capacitor.
- (b) The *very approximate* inductance of a pair of leads is approximately $1\text{ }\mu\text{H}/\text{meter}$ of lead length. For a total lead length of one inch, calculate the self-resonant frequency of the capacitor for $C = 1\text{ }\mu\text{F}$ and for $C = 1000\text{ }\mu\text{F}$.
- (c) Sketch a Bode (log-log) plot of a capacitor impedance for $C = 1000\text{ }\mu\text{F}$ and one inch leads, assuming an equivalent series resistance $\text{ESR} = 10\text{ m}\Omega$.
- (d) Why should you keep lead lengths short ?

Problem 15.2

A PC board trace is 0.1" wide, 6" long and is 0.06" above a ground plane. Estimate the trace inductance and trace resistance at room temperature. Assume that the trace is comprised of 1-ounce copper.

Problem 15.3

A microstrip line is built with two copper strips each a centimeter wide separated by a 0.005" thick kapton insulator. Estimate the inductance of a 6-inch long piece of microstrip.

Problem 15.4

A PC-board mounted inductor has an ideal value of $L = 1\text{ }\mu\text{H}$. However, the measured series resistance of this device is 0.5Ω , and a parallel-resonant frequency of 15 MHz is measured using an impedance analyzer. Generate an appropriate lumped-circuit model for this inductor.

Problem 15.5

A film capacitor has an ideal value of $2.2\text{ }\mu\text{F}$. You initially charge the capacitor to 250V, then disconnect the charging power supply and, using an oscilloscope, discover that the capacitor

voltage discharges to 100V in 60 seconds. Find a lumped circuit model of this capacitor, utilizing these findings.

Problem 15.6

A MOSFET gate driver IC has a very low impedance, and drives a MOSFET gate where the input impedance may be modeled as a capacitance of 1000 pF. The gate driver is 2 inches away on a PC board, resulting in an approximate series inductance driving the gate from the driver of 50 nH. Draw the circuit model, and find the value of series resistance that you place at the output of the gate driver to achieve critical damping of the gate circuit. Simulate your circuit and find the 10–90% risetime of the gate signal.

References

- Bartoli, M., Reatti, A., and Kazimierczuk, M., “High-frequency models of ferrite core inductors,” *International Conference on Industrial Electronics, Control and Instrumentation, (IECON '94)*, September 5–9, 1994, volume 3, pp. 1670–1675.
- Brooks, D., “Fusing Currents—When Traces Melt Without a Trace,” available at <http://www.ultracad.com>, printed in *Printed Circuit Design*, vol. 15, no. 12, Dec. 1998, pp. 53.
- Cao, Y., Groves, R., Huang, X., Zamdmer, N., Plouchart, J., Wachnik, R., King, T., and Hu, C., “Frequency-independent equivalent-circuit model for on-chip spiral inductors,” *IEEE Journal of Solid-State Circuits*, vol. 38, no. 3, March 2003, pp. 419–426.
- Demurie, S. N., and DeMey, G., “Parasitic capacitance effects of planar resistors,” *IEEE Transactions on Components, Packaging and Manufacturing Technology, Part A*, vol. 12, no. 3, September 1989, pp. 348–351.
- Dolan, J. E., and Bolton, H. R., “Capacitor ESR measurement technique,” *Eighth IEEE Pulsed Power Conference 1991*, June 16–19, 1991, pp. 228–231.
- Franco, S., “Polypropylene capacitors for snubber applications,” *Proceedings of the Thirty-First IAS Annual Meeting (IAS '96)*, October 6–10, 1996, pp. 1337–1342.
- Galbraith, J., “Reliable Precision Wirewound Resistor Design,” *IRE Transactions on Components Parts*, vol. 3, no. 3, December 1956, pp. 116–119.
- Grover, F. W., *Inductance Calculations: Working Formulas and Tables*, Dover Publications, Inc., New York, 1946.
- Jutty, M. K., Swaminathan, V., and Kazimierczuk, M. K., “Frequency characteristics of ferrite core inductors,” *Proceedings of the Electrical Electronics Insulation Conference and Electrical Manufacturing & Coil Winding Conference, 1993*, October 4–7, 1993, pp. 369–372.
- Madou A., and Martens, L., “Electrical behavior of decoupling capacitors embedded in multilayered PCBs,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 43, no. 4, November 2001, pp. 549–566.
- Manka, W., “Alternative Methods for Determining Chip Inductor Parameters,” *IEEE Transactions on Parts, Hybrids, and Packaging*, vol. 13, no. 4, December 1977, pp. 378–385.
- Massarini, A., and Kazimierczuk, M. K., “Self-capacitance of inductors,” *IEEE Transactions on Power Electronics*, vol. 12, no. 4, July 1997, pp. 671–676.

- Naishadharn, K., "Experimental equivalent-circuit modeling of SMD inductors for printed circuit applications," *IEEE Transactions on Electromagnetic Compatibility*, Volume: 43, no. 4, Nov. 2001, pp. 557–565.
- Neugebauer, T. C., Phinney, J. W., and Perreault, D. J., "Filters and Components With Inductance Cancellation," *IEEE Transactions on Industry Applications*, vol. 40, no. 2, March–April 2004, pp. 483–491.
- Reed, E. K., "Tantalum chip capacitor reliability in high surge and ripple current applications," *1994 Electronic Components and Technology Conference*, 1994, May 1–4, 1994, pp. 861–868.
- Sakabe, Y., Hayashi, M., Ozaki, T., and Canner, J. P., "High frequency measurement of multilayer ceramic capacitors," *IEEE Transactions on Components, Packaging and Manufacturing Technology, Part B: Advanced Packaging*, vol. 19, no. 1, February 1996.
- Sarjeant, W. J., Zirnheld, J., and MacDougall, F. W., "Capacitors," *IEEE Transactions on Plasma Science*, vol. 26, no. 5, October 1998, pp. 1368–1392.
- Smith, L. D., and Hockanson, D., "Distributed SPICE circuit model for ceramic capacitors," *Proceedings of the 2001 Electronic Components and Technology Conference*, May 29–June 1, 2001, pp. 523–528.
- Stroud, J., "Equivalent series resistance-the fourth parameter for tantalum capacitors," *Proceedings of the 1990 Electronic Components and Technology Conference*, May 20–23, 1990, pp. 1009–1012.
- Thompson, Marc, and Schlecht, Martin, "High Power Laser Diode Driver Based on Power Converter Technology," *IEEE Transactions on Power Electronics*, vol. 12, no. 1, January 1997, pp. 46–52.
- Ulrich, R. K., Brown, W. D., Ang, S. S., Barlow, F. D., Elshabini, A., Lenihan, T. G., Naseem, H. A., Nelms, D. M., Parkerson, J., Schaper, L. W., and Morcan, G., "Getting aggressive with passive devices," *IEEE Circuits and Devices Magazine*, vol. 16, no. 5, September 2000, pp. 16–25.
- Venkataramanan, G., "Characterization of capacitors for power circuit decoupling applications," *Industry Applications Conference, 1998*, vol. 2, October 12–15, 1998, pp. 1142–1148.
- Wadell, B. C., "Modeling circuit parasitics 1," *IEEE Instrumentation & Measurement Magazine*, vol. 1, no. 1, March 1998, pp. 31–33.
- , "Modeling circuit parasitics 2," *IEEE Instrumentation & Measurement Magazine*, vol. 1, no. 2, June 1998, pp. 6–8.
- , "Modeling circuit parasitics 3," *IEEE Instrumentation & Measurement Magazine*, vol. 1, no. 3, September 1998, pp. 28–31.
- , "Modeling circuit parasitics 4," *IEEE Instrumentation & Measurement Magazine*, vol. 1, no. 4, December 1998, pp. 36–38.
- Yu, Q., and Holmes, T. W., "A study on stray capacitance modeling of inductors by using the finite element method," *IEEE Transactions on Electromagnetic Compatibility*, vol. 43, no. 1, February 2001, pp. 88–93.

Other Useful Design Techniques and Loose Ends

In This Chapter

- Here at the end lies a potpourri of (hopefully) useful design techniques.
-

Thermal Circuits

All electrical engineers know Ohm's law, $V = IR$: that a current “flows” through a resistor, and that the current is forced by an electrical “pressure,” or voltage. The constant of proportionality between voltage and current is the circuit resistance R , in ohms. Given that we all accept Ohm's law, analogies may be made with other physical systems in an attempt to generate simple models for complicated physical processes.

With a little thought, the flow of current in a resistor may be compared to flow of water through a pipe, where current is analogous to water flow rate, and voltage is analogous to the pressure differential across the pipe that forces the water to flow. The resistance to flow is dependent on the diameter of the pipe, its length, the viscosity of the fluid, and other parameters.

With a little more thought, heat flow can be modeled with simple circuit analogies. A warm body may be heated or cooled by three mechanisms: conduction, convection, and radiation. Conduction is transfer of energy when heat is transferred by a solid; for instance, heat (in watts) is transferred down a water pipe when heat is transferred down the pipe body from a warm indoors to a cold outdoors. Convection is heat transfer due to moving fluid or air (for instance, blowing on a hot bowl of soup to cool it off). Radiation is an electromagnetic effect, and nonlinear with temperature,¹ and therefore is not simply modeled by lumped models.

Using relatively simple models of static and time-varying heat transfer, temperature effects can be modeled. These techniques are useful for heat sink design, PC thermal design, and for gaining a basic understanding of heat transfer processes.

¹ Radiation from a warm body is proportional to T^4 , where T is temperature in Kelvin. In many instances, for instance at room temperature, radiation can be neglected compared to conductive and convective heat transfer. But, for forced-air convection (for instance, rapid air flow over the fins of a heat sink) the *convective* heat transfer may be the dominant effect.

Steady-State Model of Conductive Heat Transfer

In many cases of interest, where there is no air flow and near room temperature, conduction will be the dominant heat transfer effect. For instance, inside an IC package, there is no air flow, yet heat is transferred from the electrical junctions where the heat is being generated, out to the external world (for instance, to a heat sink or the PC board). In order to relate this to a simple circuit model, consider a simple resistor circuit, and its thermal circuit analogy (**Figure 16-1**).



Figure 16-1: Analogy between current flow in a resistor (a) and one-dimensional heat flow through a long, skinny solid (b).

If we examine the electrical case, current (I , in amperes) flows, and is forced by a voltage differential $V_2 - V_1$; current flows from the higher voltage to the lower voltage, and is given by:

$$I = \frac{V_2 - V_1}{R_{elec}} \quad [16-1]$$

The constant of proportionality between current and voltage is the electrical resistance in ohms (R_{elec}). The electrical resistance is given by:

$$R_{elec} = \frac{l}{\sigma A} \quad [16-2]$$

where l is the length of the resistor, A is the cross-sectional area through which the current flows, and σ is the electrical conductivity of the material that makes up the resistor, in $\Omega^{-1}\text{m}^{-1}$. Current flows easier if the electrical resistor is shorter and fatter.

Next, let's consider the thermal case. Consider a long, skinny rod of material, which is made of a good heat conductor such as a metal. It is in good thermal contact with two heat sinks, one at either end of the rod. We assume that there is a temperature difference between the two ends of the rod. For this simple case, it may be approximated that heat only flows in one direction. Heat (P , in watts) flows, and is forced by a temperature differential $T_2 - T_1$; a constant power flows from the higher temperature to the lower temperature, and is given by:

$$P = \frac{T_2 - T_1}{R_{TH}} \quad [16-3]$$

The constant of proportionality between heat flow and temperature is the thermal resistance R_{TH} , which has units of degrees per watt. This is the proportionality constant which tells you how many degrees you need to put across an element to cause one Watt of heat to flow. Most metals, like copper and aluminum, have a high thermal conductivity (that is, they transfer heat very well).

Analogous to electrical resistance, thermal resistance is given by:

$$R_{TH} = \frac{l}{kA} \quad [16-4]$$

where l is the length of the material through which heat flows, A is the cross-sectional area through which the heat flows, and k is the thermal conductivity of the material that makes up the resistor, in watts/(meter-degree). As in the electrical case, thermal resistance is lower if the conducting path is shorter and has more area.

Still air, which is a good insulator, has a thermal conductivity $k \approx 0.03$ watt/meter-°C. Copper, which is a good conductor of heat, has $k \approx 400$ watt/meter-°C. These numbers illustrate why styrofoam, which is full of trapped air, is a good insulator, while copper and aluminum are not. This is one reason why metals aren't often used as thermal insulators!

Thermal Energy Storage

The effects of energy storage in thermal systems are also demonstrated by analogy. Just as energy is stored in a capacitor as charge on the capacitor plates, energy is stored in the mechanical structure of a mass, in the vibration of the atoms. A body at a higher temperature has a higher amount of stored energy than the same body at a lower temperature.

In the electrical case, when a linear capacitor is charged from an initial voltage (V_i) to a final voltage (V_f), charge (q , in coulombs) is stored on the plates of the capacitor, as:

$$q = C(V_f - V_i) \quad [16-5]$$

where C is electrical capacitance, in coulombs per volt. During charging, a current flows into the capacitor, as:

$$i = \frac{dq}{dt} = C \frac{dv(t)}{dt} \quad [16-6]$$

The charge that flows during the charging interval is stored on the plates of the capacitor.

In the thermal case, the thermal capacitance of a mass determines how well the mass stores energy; the units of heat capacity are joules per degree. The thermal capacitance is given by:

$$C_{TH} = Mc_v = V\rho c_v \quad [16-7]$$

where ρ is the density of the material (kg/m³), M is the total mass of the material that is being heated or cooled (kg), V is the volume of the material (m³) that is being heated or cooled, and c_v is a material property called specific heat which is a material property (joules/kg-degree).

When a mass is heated from an initial temperature (T_i) to a final temperature (T_f), energy (E , in joules, or watt-seconds) must be added to the mass to heat it up, as:

$$E = C_{TH}(T_f - T_i) \quad [16-8]$$

where C_{TH} is the heat capacity, in joules per degree. This is analogous to storage of charge in an electrical capacitor. During the heating process, a power (watts) flows into the mass, as:

$$P(t) = \frac{dE}{dt} = C_{TH} \frac{dT}{dt} \quad [16-9]$$

where dT is the temperature change in a time interval dt . The energy that flows during the heating interval is stored in the mass of the body being heated.

In order to use the simple transient model, consider the arrangement shown in **Figure 16-2**. The block is a block of material, which is being heated by a source of energy (perhaps a blow torch) that delivers P_o watts of power to the mass being heated.

The big “wall” that the block is attached to is a heat sink at ambient temperature T_A . It is assumed that the heat sink is large and sufficiently cooled so that its temperature does not change during the duration of the test. What is the temperature profile of the block of material after the heat source is applied?

Using the circuit analogies, the heater is a source of power $P(t)$; this is modeled as a step in applied heat (analogous to a circuit current) with amplitude P_o (watts). Circuit “ground” is the ambient temperature T_A . Heat is conducted away from the block into the heat sink through a thermal resistance R_{TH} . Simultaneously, energy is stored in the thermal capacitance of the block C_{TH} .

The heat source is modeled as a “step” of power (**Figure 16-3**). When the heat is first turned on (at time $t = 0$), the block is at ambient temperature, or the same temperature of the heat sink, T_A . The temperature of the block $T_B(t)$ increases as:

$$T_B(t) = P_o R_{TH} \left(1 - e^{-\frac{t}{\tau_{TH}}} \right) \quad [16-10]$$

The temperature of the block reaches the final temperature with the RC thermal time constant (τ_{TH}), just as in the step response of a current-source driven RC electrical circuit. The circuit will continue “charging” (i.e., the block’s temperature will increase) until the input power is exactly balanced by the heat conducted away through R_{TH} to the heat sink. After a few time constants, the block approaches the final temperature:

$$T_F = P_o R_{TH} \quad [16-11]$$

Figure 16-2: Use of simple transient model to find a crude estimate of block temperature vs. time. (a) Physical arrangement showing a blowtorch delivering power $P(t)$ to a block which is heated up. The temperature at the center of the block is T_B and the block is mounted to a heat sink that operates at temperature T_A . (b) Circuit model showing thermal resistance R_{TH} and thermal capacitance C_{TH} .

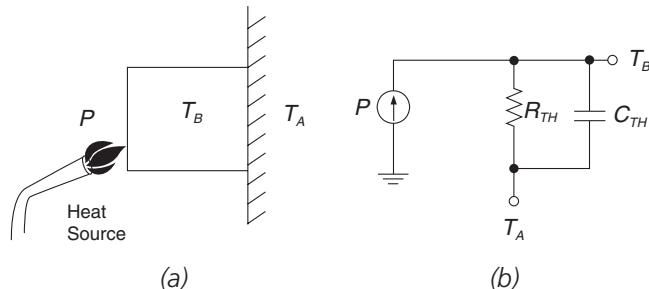
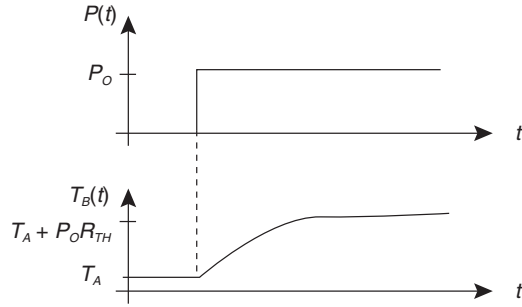


Figure 16-3: Use of simple transient model. Input “step” of power $P(t)$ and temperature of block $T_B(t)$.



Now, to be sure, some of the details of heat transfer are glossed over using this simple circuit; for instance, there is some heat transferred between the block and the surrounding air. However, if the heat sink is well-designed, there will be much more efficient transfer of heat to the heat sink than to the air (after all, that is what a heat sink is for). Secondly, the lumped circuit model is valid only for “low frequencies,” or for temperatures that change on a time scale much longer than the thermal time constant. That means that the temperature throughout the body changes the same everywhere. And, the specific heat and thermal conductivity of materials vary with temperature, but over a limited temperature range they can be approximated as being constant.

For engineering calculations, and if you check your approximations, these calculations give you very useful results. You can always do a finite element analysis to get more exact results.

To summarize: Using relatively simple models, static and dynamic heat transfer can be modeled with simple circuit analogies. Important questions may be answered, such as:

- When you apply a source of heat to a material, what is the total temperature rise in the material?
- What do those specifications in manufacturers’ data sheets mean for junction-to-pin and junction-to-ambient thermal resistance?
- On what time scale does the temperature change when a material is heated and how fast does the temperature reach its final value?

The electrical-to-thermal analogies are summarized in **Table 16-1**.

Table 16-1: Summary of electrical and thermal analogies

	Electrical system	Thermal system
Stored quantity	charge q	energy E
“Flow” quantity	current I	power P
Pressure which forces flow	voltage V	temperature T
Resistance to flow	$R = \frac{1}{\sigma A}$	$R_{TH} = \frac{l}{kA}$
Capacitance	$C = \varepsilon \frac{A}{d}$	$C_{TH} = \rho c_v V$
Time constant	$\tau_{elec} = RC$	$\tau_{TH} = R_{TH} C_{TH}$

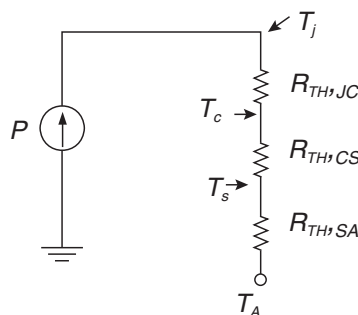
Using Thermal Circuit Analogies to Determine Static Semiconductor Junction Temperature

The thermal model of a semiconductor mounted to a heat sink is shown in **Figure 16-4**. The transistor dissipates power, and that is indicated by current source with value P . Ambient temperature is T_A , and the heat sink surface temperature and case temperature of the semiconductor are T_s and T_c respectively. The transistor junction temperature is T_j , and our goal in designing a heat-sink system is to guarantee that the junction temperature doesn't exceed a safe level.

The thermal resistances for heat conduction through the transistor to the heat sink surface are $R_{TH,JC}$ and $R_{TH,CS}$. $R_{TH,JC}$ is the thermal resistance from the junction of the transistor (where the power is dissipated) to the case of the transistor. The thermal resistance $R_{TH,CS}$ is thermal resistance from the case of the transistor to the heat sink. This value of thermal resistance depends on the contact area, how well you torque the transistor down to the heat sink, and the type of thermal interface² material that you use.

The thermal resistance from the heat sink to ambient air ($R_{TH,SA}$) is a function of the heat sink area and whether you cool the heat sink with forced air or not. Heat sink manufacturers will specify this number for a given heat sink area and air flow.

Figure 16-4: Thermal model of a semiconductor mounted to a heat sink.



Mechanical Circuit Analogies

In many instances, solutions for natural frequency and mode shapes of electrical circuits can be found by considering behavior of analogous mechanical systems. Operation of the mechanical system, consisting of masses (m), dampers (c) and springs (k) may be easy to visualize by using simple physical reasoning. The solution can then be mapped to an analogous system consisting of inductors (L), resistors (R) and capacitors (C).

In the vibrating mechanical system, energy is transferred back and forth between kinetic energy in the mass and potential energy stored in the springs. Equivalently, in the electromagnetic system, energy is sloshed back and forth between magnetic energy stored in inductors and electrical energy stored in the capacitors. A simple example explains this duality.

² For example, do you use thermal grease, or a mica pad, etc.

Mechanical system:

The mechanical system for this example consists of two masses on frictionless rollers, connected by a massless spring. The force exerted by the spring is kx , where k is the mechanical spring constant in Newtons/meter. The two state variables for this system are the horizontal positions of the two masses, x_1 and x_2 , defined with respect to a fixed position on the Earth.

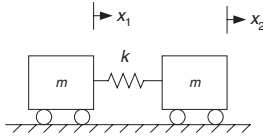


Figure 16-5: Mechanical two degree-of-freedom circuit.

Considering the first mass on the left, Newton's law gives³:

$$m\ddot{x}_1 = k(x_2 - x_1)$$

Applying the same reasoning to the right-hand mass results in the coupled equations of motion for the two masses.

$$m\ddot{x}_1 + k(x_1 - x_2) = 0$$

$$m\ddot{x}_2 + k(x_2 - x_1) = 0$$

Assuming sinusoidal variations in position, with $x = X_o e^{j\omega t}$

$$-\omega^2 \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} + \begin{bmatrix} k & -k \\ -k & k \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} = 0$$

or:

$$\begin{bmatrix} -\omega^2 m + k & -k \\ -k & -\omega^2 m + k \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} = 0$$

The natural frequencies are found by solving the determinant of the above matrix. Solutions for this set of simultaneous equation result in both two natural frequencies, (ω) and mode shapes for

Electrical system:

The electrical dual for this system consists of an L-C-L circuit. This is the equivalent circuit of a section of a lossless transmission line. The two state variables for this system are the loop currents i_1 and i_2 , defined as shown.

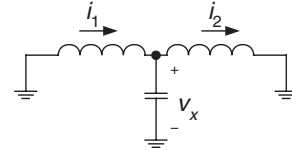


Figure 16-8: Electrical two degree-of-freedom circuit

Considering the first inductor on the left:

$$L\dot{i}_1 = -v_x$$

Taking the derivative results in:

$$L\ddot{i}_1 = -\frac{dv_x}{dt} = \frac{-(i_1 - i_2)}{C}$$

Applying the same reasoning to the right-hand inductor results in the coupled equations:

$$L\ddot{i}_1 + \frac{1}{C}(i_1 - i_2) = 0$$

$$L\ddot{i}_2 + \frac{1}{C}(i_2 - i_1) = 0$$

Assuming sinusoidal variations in current, with $i = I_o e^{j\omega t}$

$$-\omega^2 \begin{bmatrix} L & 0 \\ 0 & L \end{bmatrix} \begin{Bmatrix} i_1 \\ i_2 \end{Bmatrix} + \begin{bmatrix} \frac{1}{C} & -\frac{1}{C} \\ -\frac{1}{C} & \frac{1}{C} \end{bmatrix} \begin{Bmatrix} i_1 \\ i_2 \end{Bmatrix} = 0$$

or:

$$\begin{bmatrix} -\omega^2 L + \frac{1}{C} & -\frac{1}{C} \\ -\frac{1}{C} & -\omega^2 L + \frac{1}{C} \end{bmatrix} \begin{Bmatrix} i_1 \\ i_2 \end{Bmatrix} = 0$$

³ Remember the notation: \ddot{x} is the second derivative with respect to time of the variable x . \dot{x} is the first derivative.

position of the masses (the x_1, x_2 vector). Solving for allowable natural frequencies results in:

Natural Frequency	Mode Shape
$\omega_a = 0$	$\{x\} = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$
$\omega_b = \sqrt{\frac{2k}{m}}$	$\{x\} = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}$

Now, the above set of natural frequencies and mode shapes can be easily found by inspection. For instance, the first mode shape corresponds to the case where both masses move to the right (or left, depending on the initial conditions) in unison. In this case, there is no stretching of the spring, and no vibration. The spring could be replaced by a massless rigid rod without disturbing the movement of the masses, which are both in simple translation. Therefore, the natural frequency = 0.

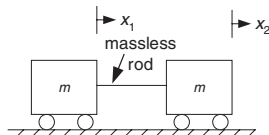


Figure 16-6: First (lower frequency mode).

For the second mode, both masses vibrate, but 180° out of phase. Therefore, the middle of the spring doesn't move (i.e., this position is a node). By symmetry, we could put a brick wall there and break up the single spring into two springs of spring constant $2k$, and solve for the vibration of each half-circuit independently.

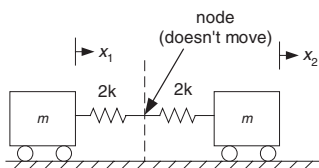


Figure 16-7: Second (higher frequency mode).

Solutions for this set of simultaneous equation result in both two natural frequencies, (ω) and mode shapes for position of the masses (the i_1, i_2 vector). Solving for allowable natural frequencies results in:

Natural Frequency	Mode Shape
$\omega_a = 0$	$\{i\} = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$
$\omega_b = \sqrt{\frac{2}{LC}}$	$\{i\} = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}$

Now, the above set of natural frequencies and mode shapes can be easily found by inspection. For instance, the first mode shape corresponds to the case where both loop currents flow clockwise (or counterclockwise, depending on the initial conditions). In this case, there is no current flow in the capacitor, and no vibration. In fact, for this mode of operation, the capacitor can be removed from the circuit since there is no current flow in it. Therefore, the natural frequency = 0.

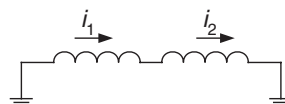


Figure 16-9: First (lower frequency mode).

For the second mode, there is oscillation in i_1 and i_2 , but 180° out of phase. We can model the two half-circuits by breaking up the capacitor into two capacitors, each of value $C/2$. By symmetry, there is no current flow across the boundary. Therefore, the two half-circuits each vibrate independently.

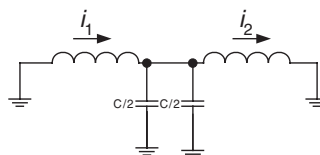


Figure 16-10: Second (higher frequency mode).

These same techniques can be used for systems with different boundary conditions, and for systems with more degrees-of-freedom.

Example 16.1: Using mechanical circuit analogies

As an example, consider the circuit in **Figure 16-11a**. The expected natural frequency for the first mode of oscillation is zero, corresponding to the case when inductor currents i_1 and i_2 are equal value and in the same direction. For the first mode, the capacitor current is forever zero, and hence the capacitor voltage remains constant. Therefore, there is no sinusoidal oscillation for the first mode.

In the second mode of oscillation i_1 and i_2 are in opposite directions and of equal value, and we refer to the circuits of **Figure 16-11b** and **Figure 16-11c**. The second mode of oscillation has an expected natural frequency of:

$$\omega_{o, \text{mode2}} = \sqrt{\frac{2}{LC}} = 1.41 \text{ rad/sec} \quad [16-12]$$

equivalent to a natural frequency of 0.225 Hz. The PSPICE output (**Figure 16-12**⁴) shows the inductor currents with initial conditions set $i_1 = 1$ and $i_2 = -1$. As expected, the circuit oscillates at 0.225 Hz.

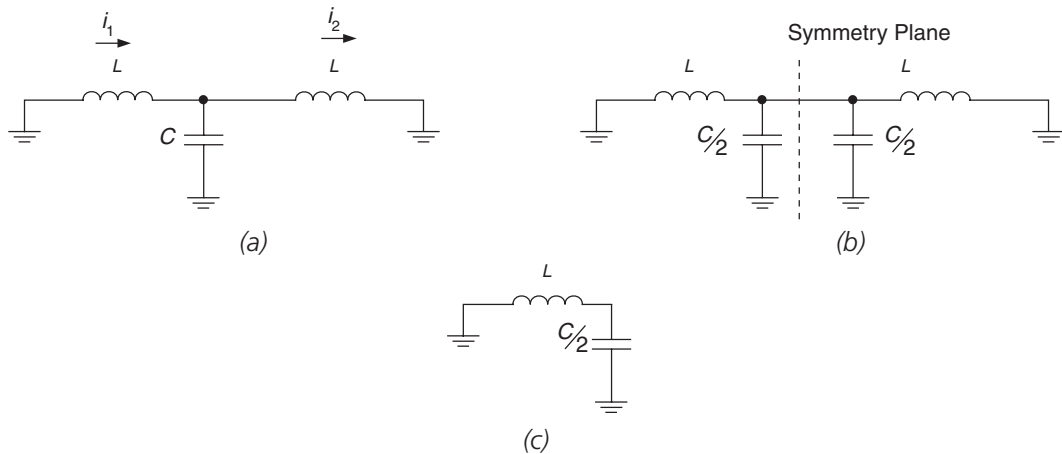


Figure 16-11: Circuit for example 16.1. (a) Original circuit. (b) Original circuit, redrawn showing the symmetry plane. For the second mode of oscillation, there is no current across the symmetry plane. (c) Simplified circuit for the second mode of oscillation, exploiting symmetry.

⁴ The small resistor is added because SPICE has convergence problems if there are voltage loops containing inductors.

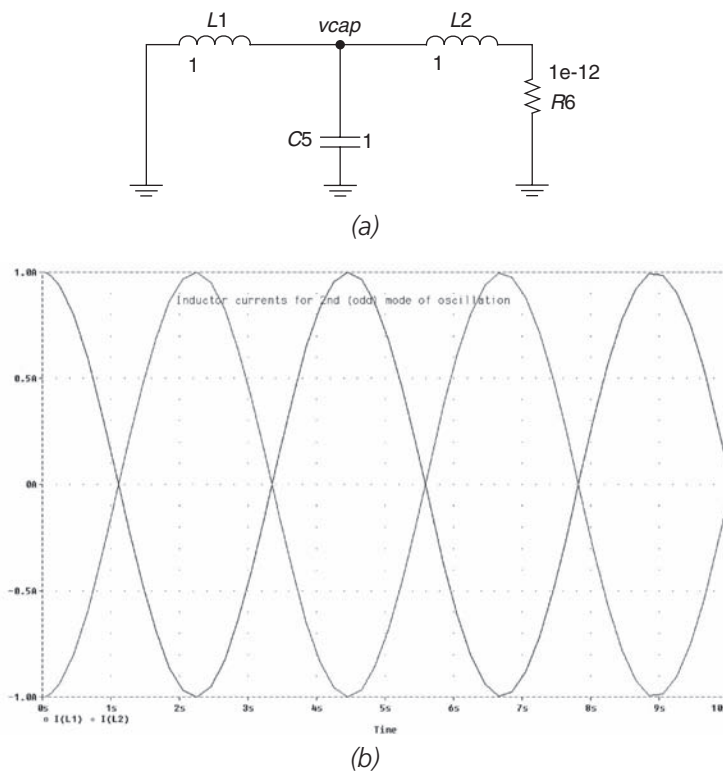


Figure 16-12: Second mode of oscillation. (a) PSPICE circuit. (b) PSPICE result for second mode, showing inductor currents of equal magnitude and 180° out of phase.

For the second mode of operation, there are two independent circuits, each with inductance L and capacitance $C/2$. Therefore, the characteristic impedance of each circuit is:

$$Z_o = \sqrt{\frac{L}{C/2}} = \sqrt{\frac{2L}{C}} = 1.41\Omega \quad [16-13]$$

Therefore, we expect the peak-to-peak amplitude of oscillation of the capacitor voltage to be $2\sqrt{2}V$, as shown in **Figure 16-13**.

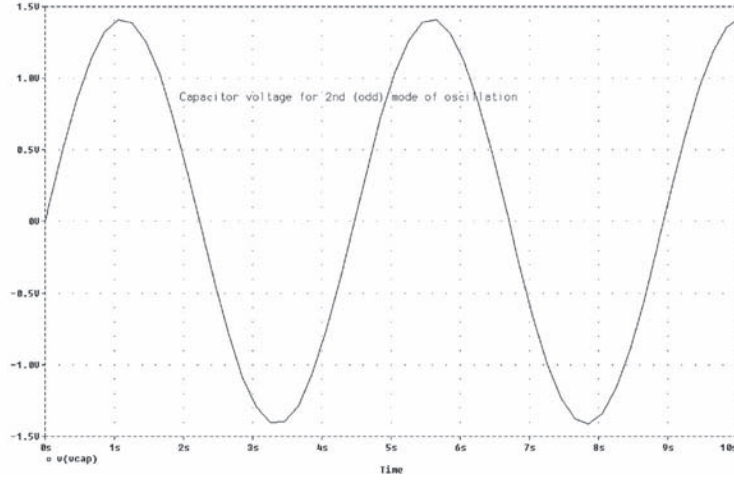


Figure 16-13: Capacitor voltage for the second mode of oscillation.

The Translinear Principle

Translinear⁵ circuits exploit the well-known exponential relationship between transistor base-emitter voltage and collector current, a relationship that holds over many orders of magnitude of collector current. Consider the translinear circuit of **Figure 16-14**. We have a loop of V_{BE} s as indicated with the arrows. If we do KVL around the V_{BE} loop we get:

$$-V_{BE1} + V_{BE2} - V_{BE3} + V_{BE4} = 0 \quad [16-14]$$

Let's recall the logarithmic expression for V_{BE} of a transistor operated in the forward-active region:

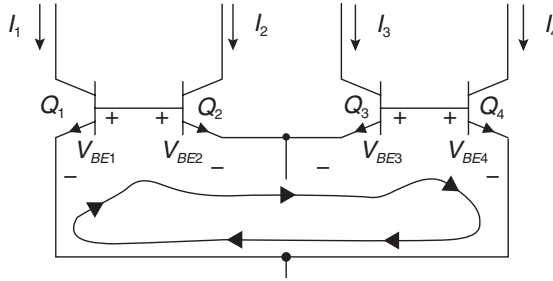
$$V_{BE} \approx \frac{kT}{q} \ln \left(\frac{I_C}{I_S} \right) \quad [16-15]$$

where I_S is the reverse saturation current of the transistor. In the translinear circuit above, if all transistors are identical, we can write:

$$-\frac{kT}{q} \ln \left(\frac{I_{C1}}{I_S} \right) + \frac{kT}{q} \ln \left(\frac{I_{C2}}{I_S} \right) - \frac{kT}{q} \ln \left(\frac{I_{C3}}{I_S} \right) + \frac{kT}{q} \ln \left(\frac{I_{C4}}{I_S} \right) = 0 \quad [16-16]$$

⁵ The term “translinear” was coined by Barrie Gilbert.

Figure 16-14: Translinear circuit, showing a loop of V_{BE} s.



The solution for this is:

$$I_{C2}I_{C4} = I_{C1}I_{C3} \quad [16-17]$$

For a circuit of this type with a loop of V_{BE} s and identical transistors, we can state the translinear principle:

The product of the clockwise currents equals the product of the counter-clockwise currents, or:

$$\prod \left. \frac{I_C}{I_S} \right|_{cw} = \prod \left. \frac{I_C}{I_S} \right|_{ccw} \quad [16-18]$$

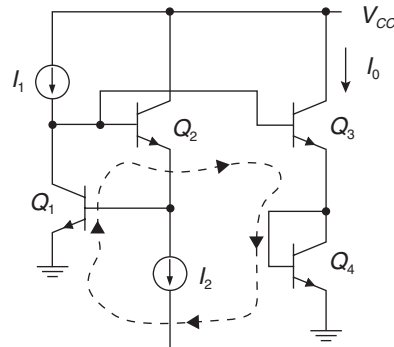
Let's apply the translinear principle to another translinear circuit (**Figure 16-15**). This circuit has a loop of V_{BE} s including Q_1 , Q_2 , Q_3 and Q_4 . Going around the loop, we see that I_{C1} and I_{C2} are counter-clockwise currents, and I_{C3} and I_{C4} are clockwise currents. Therefore, for this circuit:

$$I_{C1}I_{C2} = I_{C3}I_{C4} \quad [16-19]$$

This means we can express the output current I_o as:

$$I_o = \sqrt{I_1 I_2} \quad [16-20]$$

Figure 16-15: Another translinear circuit that performs a square-root function. The dotted line indicates the loop of V_{BE} s over which the translinear principle is used.



Input Impedance of Infinitely Long Resistive Ladder

Let's find the input impedance of an infinitely long ladder with series element of impedance Z and shunt element with admittance Y (**Figure 16-16**). Since this ladder is infinitely long, we can say that the impedance at position 1 is the input impedance at position 2. Mathematically, we can express this as:

$$Z_{in} = Z + \frac{1}{Y} \parallel Z_{in} = Z + \frac{YZ_{in}}{Y + Z_{in}} \quad [16-21]$$

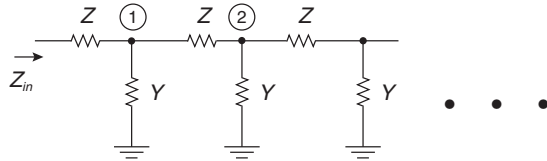
We can rearrange this to get a quadratic equation for Z_{in} :

$$Z_{in}^2 - ZZ_{in} - ZY = 0 \quad [16-22]$$

The solution to this equation is:⁶

$$Z_{in} = \frac{Z}{2} \left(1 + \sqrt{1 + \frac{4}{ZY}} \right) \quad [16-23]$$

Figure 16-16: Infinitely long resistive ladder.

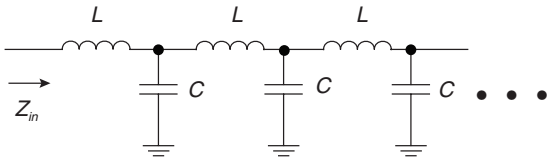


Transmission Lines 101

The interesting result from the infinitely long resistive ladder can be used to find the input impedance of the ideal transmission line (**Figure 16-17**). We recognize that using the formulation above:

$$\begin{aligned} Z &= j\omega L \\ Y &= j\omega C \end{aligned} \quad [16-24]$$

Figure 16-17: Ideal transmission line.



This means that the solution for the input impedance Z_{in} of the ideal transmission line is:

$$Z_{in} = \frac{j\omega L}{2} \left(1 + \sqrt{1 + \frac{4}{(j\omega L)(j\omega C)}} \right) \quad [16-25]$$

⁶ We've thrown away one of the solutions to this quadratic which has no physical meaning.

Chapter 16

For an ideal transmission line, there are lots of L and C lumps, so each lump is very small. We can express this as L and $C \rightarrow 0$. This means we can find the input impedance in this limit as:

$$Z_{in} = \frac{j\omega L}{2} \left(\sqrt{\frac{4}{(j\omega L)(j\omega C)}} \right) = \sqrt{\frac{L}{C}} \equiv Z_o \quad [16-26]$$

This characteristic impedance Z_o tells us the ratio of voltage to current along the line.

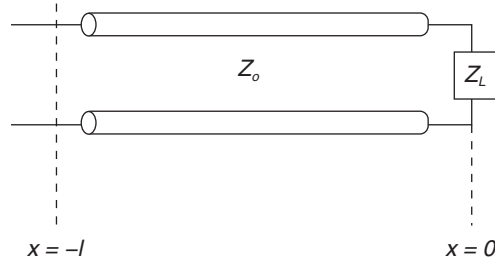
Finding the input impedance of a finite-length transmission line

A transmission line can be modeled as a lumped circuit (i.e., with inductors and capacitors) if the length of the line is much smaller than a wavelength. Let's consider a transmission line of characteristic impedance Z_o , length l , and terminated by an impedance Z_L at $x = 0$, as shown in **Figure 16-18**. The basic transmission line equations are:

$$\begin{aligned} V(x, \omega) &= V_+ e^{\frac{-j\omega x}{c}} + V_- e^{\frac{j\omega x}{c}} \\ I(x, \omega) &= \frac{V_+}{Z_o} e^{\frac{-j\omega x}{c}} - \frac{V_-}{Z_o} e^{\frac{j\omega x}{c}} \end{aligned} \quad [16-27]$$

where V_+ is the forward-traveling wave on the line, V_- is the reflected wave, and c is the speed of wave propagation down the line.

Figure 16-18: Unmatched transmission line. The line has characteristic impedance Z_o and is terminated with impedance Z_L .



The reflection coefficient is the ratio of the reflected wave amplitude to the forward wave amplitude, or:

$$\Gamma = \frac{V_-}{V_+} \quad [16-28]$$

We can use this in the basic transmission line equation to find:

$$\begin{aligned} V(x, \omega) &= V_+ \left(e^{\frac{-j\omega x}{c}} + \Gamma e^{\frac{j\omega x}{c}} \right) \\ I(x, \omega) &= \frac{V_+}{Z_o} \left(e^{\frac{-j\omega x}{c}} - \Gamma e^{\frac{j\omega x}{c}} \right) \end{aligned} \quad [16-29]$$

At $x = 0$, we have the boundary condition that $V/I = Z_L$, or:

$$Z_o \left(\frac{1 + \Gamma}{1 - \Gamma} \right) = Z_L \quad [16-30]$$

This enables us to solve for the reflection coefficient as:

$$\Gamma = \frac{Z_L - Z_o}{Z_L + Z_o} \quad [16-31]$$

This result makes sense, because we know that on a matched transmission line (with $Z_L = Z_o$) there are no reflections from the end of the line. With an unmatched line ($Z_L \neq Z_o$) there are reflections. We can now use the reflection coefficient result to help solve for the impedance at every point along the line:

$$Z(x, \omega) = \frac{V(x, \omega)}{I(x, \omega)} = \frac{V_+ \left(e^{\frac{-j\omega x}{c}} + \Gamma e^{\frac{j\omega x}{c}} \right)}{\frac{V_+}{Z_o} \left(e^{\frac{-j\omega x}{c}} - \Gamma e^{\frac{j\omega x}{c}} \right)} = Z_o \left(\frac{e^{\frac{-j\omega x}{c}} + \left(\frac{Z_L - Z_o}{Z_L + Z_o} \right) e^{\frac{j\omega x}{c}}}{e^{\frac{-j\omega x}{c}} - \left(\frac{Z_L - Z_o}{Z_L + Z_o} \right) e^{\frac{j\omega x}{c}}} \right) \quad [16-32]$$

By multiplying through by $Z_L + Z_o$ we get:

$$Z(x, \omega) = Z_o \frac{\left((Z_L + Z_o) e^{\frac{-j\omega x}{c}} + (Z_L - Z_o) e^{\frac{j\omega x}{c}} \right)}{\left((Z_L + Z_o) e^{\frac{-j\omega x}{c}} - (Z_L - Z_o) e^{\frac{j\omega x}{c}} \right)} \quad [16-33]$$

Next, we make use of the formulae:

$$\begin{aligned} \cos(a) &= \frac{1}{2} (e^{ja} + e^{-ja}) \\ \sin(a) &= \frac{1}{2j} (e^{ja} - e^{-ja}) \end{aligned} \quad [16-34]$$

to reduce the above equations to:

$$Z(x, \omega) = Z_o \frac{Z_L \cos\left(\frac{\omega x}{c}\right) - jZ_o \sin\left(\frac{\omega x}{c}\right)}{Z_o \cos\left(\frac{\omega x}{c}\right) - jZ_L \sin\left(\frac{\omega x}{c}\right)} \quad [16-35]$$

We're generally interested in what the input impedance is looking into the input of the transmission line (at $x = -l$). This is found by:⁷

$$Z(-l, \omega) = Z_o \frac{Z_L \cos\left(\frac{\omega l}{c}\right) + jZ_o \sin\left(\frac{\omega l}{c}\right)}{Z_o \cos\left(\frac{\omega l}{c}\right) + jZ_L \sin\left(\frac{\omega l}{c}\right)} = Z_o \frac{\frac{Z_L}{Z_o} + j \tan\left(\frac{\omega l}{c}\right)}{1 + j \frac{Z_L}{Z_o} \tan\left(\frac{\omega l}{c}\right)} \quad [16-36]$$

⁷ Remember that $\cos(-l) = \cos(l)$ and that $\sin(-l) = -\sin(l)$ since cosine is an even function and sine is an odd function. Also, remember that $\sin(x)/\cos(x) = \tan(x)$.

Chapter 16

Next, let's assume that we are operating at a frequency low enough⁸ so that $\omega l \ll c$. We can then approximate the tangent in the numerator and denominator by the Taylor series expansion:

$$\tan(x) = x + \frac{x^3}{3} + \frac{2x^5}{15} + \dots \approx x \text{ if } x \ll 1 \quad [16-37]$$

Also, remember that for small x the following holds:

$$\frac{1}{1+x} \approx 1-x \text{ if } x \ll 1 \quad [16-38]$$

So, we can further estimate the impedance looking into the line as:

$$Z(-l, \omega) \approx Z_o \frac{\frac{Z_L}{Z_o} + j\left(\frac{\omega l}{c}\right)}{1 + j\frac{Z_L}{Z_o}\left(\frac{\omega l}{c}\right)} \approx \frac{Z_L + jZ_o\left(\frac{\omega l}{c}\right)}{1 + j\frac{Z_L}{Z_o}\left(\frac{\omega l}{c}\right)} \quad [16-39]$$

Let's see what we get if we assume that the line is terminated with a low impedance, with $Z_L \ll Z_o$. In this limiting case:

$$Z(-l, \omega) \approx Z_L + j\omega\left(\frac{Z_o l}{c}\right) \approx Z_L + j\omega L_{eq} \quad [16-40]$$

This equivalent circuit is shown in **Figure 16-19a** where we see that the approximate lumped inductance is $L_{eq} \approx Z_o l / c$.

Now, in the case where $Z_L \gg Z_o$, the approximate impedance for the short line is:

$$Z(-l, \omega) \approx \frac{Z_L + jZ_o\left(\frac{\omega l}{c}\right)}{j\frac{Z_L}{Z_o}\left(\frac{\omega l}{c}\right)} \approx \frac{1}{Z_L} + \frac{1}{j\omega\left(\frac{1}{Z_o c}\right)} \approx \frac{1}{Z_L} + \frac{1}{j\omega C_{eq}} \quad [16-41]$$

The equivalent circuit for this case is shown in **Figure 16-19b** where we see an equivalent capacitance $C_{eq} \approx 1/(Z_o c)$ in parallel with the load impedance.

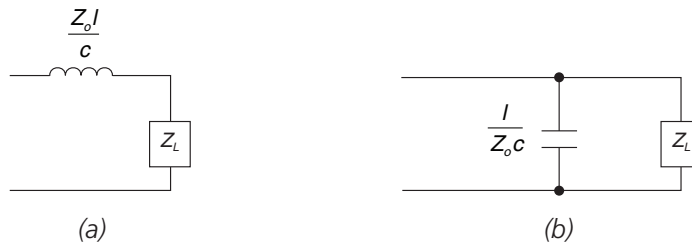


Figure 16-19: Equivalent circuit of short, unmatched transmission line.

(a) With $Z_L \ll Z_o$. (b) With $Z_L \gg Z_o$.

⁸ Or, the line is short enough.

Example 16.2: Transmission line calculation

Find the approximate equivalent circuit, valid at low frequencies for the following scenarios. Assume that the wave propagation speed $c = 3 \times 10^8$ m/s.

- 1 meter long transmission line, with characteristic impedance $Z = 75\Omega$, terminated with $Z_L = 5\Omega$.
- 1 meter long transmission line, with characteristic impedance $Z = 75\Omega$, terminated with $Z_L = 1000\Omega$.

SOLUTION

- For $Z_L = 5\Omega$, the approximate value of series inductance is:

$$L_{eq} \approx \frac{Z_o l}{c} = \frac{(75)(1)}{3 \times 10^8} \approx 250 \text{ nH} \quad [16-42]$$

- In the second scenario, the approximate value of shunt capacitance is:

$$C_{eq} \approx \frac{l}{Z_o c} = \frac{(1)}{(75)(3 \times 10^8)} \approx 44 \text{ pF} \quad [16-43]$$



Figure 16-20: Equivalent circuit of short, unmatched transmission line.

(a) With $Z_L = 5\Omega \ll Z_o$. (b) With $Z_L = 1000\Omega \gg Z_o$.

Node Equations and Cramer's Rule

Cramer's rule is a useful linear algebraic "recipe" which can help you solve any linear system of constant-coefficient equations, and hence to find the transfer function of a transistor amplifier. From a linear algebra point of view, the system of equations

$$[A]\{x\} = \{b\} \quad [16-44]$$

can be solved using Cramer's rule, where $[A]$ is an $m \times m$ matrix, $\{x\}$ is a column vector of m unknowns and $\{b\}$ is a column vector of inputs. Expanding this matrix notation results in for a system with four equations and four unknowns:

$$[A]\{x\} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \quad [16-45]$$

Using Cramer's rule, if you wanted to find unknown x_3 you solve the following:

$$x_3 = \frac{\det \begin{bmatrix} a_{11} & a_{12} & b_1 & a_{14} \\ a_{21} & a_{22} & b_2 & a_{24} \\ a_{31} & a_{32} & b_3 & a_{34} \\ a_{41} & a_{42} & b_4 & a_{44} \end{bmatrix}}{\det \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}} \quad [16-46]$$

This looks like a mess, but can be easily manipulated using simple linear algebra. The term “det” denotes the “determinant” of a matrix. For a 2×2 matrix, the determinant is found as follows:

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad [16-47]$$

For a 3×3 matrix, the determinant is:

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \quad [16-48]$$

For an $N \times N$ matrix, it's more complicated but still doable. Consult a linear algebra textbook for the recipe.

Example 16.3: Using Cramer's rule to solve simultaneous linear equations

Let's use a simple linear algebra example to show the use of Cramer's rule. We'll solve the following simultaneous linear equations for x , y and z .

$$\begin{aligned} x + y + z &= 5 \\ x - y + 3z &= -3 \\ 2x + 2y + 3z &= 10 \end{aligned} \quad [16-49]$$

Putting this into matrix form results in:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 3 \\ 2 & 2 & 3 \end{bmatrix} \begin{Bmatrix} x \\ y \\ z \end{Bmatrix} = \begin{Bmatrix} 5 \\ -3 \\ 10 \end{Bmatrix} \quad [16-50]$$

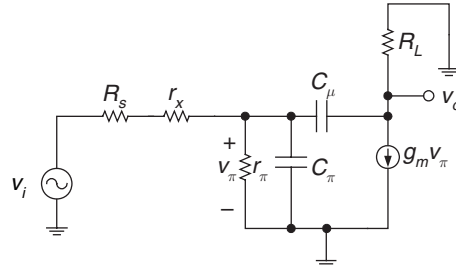
We solve for x , y and z as follows:

$$\begin{aligned}
 x &= \frac{\det \begin{bmatrix} 5 & 1 & 1 \\ -3 & -1 & 3 \\ 10 & 2 & 3 \end{bmatrix}}{\det \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 3 \\ 2 & 2 & 3 \end{bmatrix}} = \frac{(5)(-3-6)-1(-9-30)+1(-6+10)}{1(-3-6)-1(3-6)+1(2+2)} = \frac{-2}{-2} = 1 \\
 y &= \frac{\det \begin{bmatrix} 1 & 5 & 1 \\ 1 & -3 & 3 \\ 2 & 10 & 3 \end{bmatrix}}{\det \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 3 \\ 2 & 2 & 3 \end{bmatrix}} = \frac{(1)(-9-30)-(5)(3-6)+(1)(10+6)}{-2} = \frac{-8}{-2} = 4 \quad [16-51] \\
 z &= \frac{\det \begin{bmatrix} 1 & 1 & 5 \\ 1 & -1 & -3 \\ 2 & 2 & 10 \end{bmatrix}}{\det \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 3 \\ 2 & 2 & 3 \end{bmatrix}} = \frac{(1)(-10+6)-(1)(10+6)+(5)(2+2)}{-2} = \frac{0}{-2} = 0
 \end{aligned}$$

By direct substitution, we find that the values $x = 1$, $y = 4$ and $z = 0$ are indeed all correct.

Another simple example illustrates the use of Cramer's rule in the circuit realm. Shown in **Figure 16-21** is the small-signal model for a common emitter amplifier. Let's generate the node equations and solve for the transfer function using Cramer's rule.

Figure 16-21: Small-signal model of common-emitter amplifier.



We'll lump R_s and r_x together in a single effective source resistance R_s' , with $G_s' = 1/R_s'$. At node v_π , we sum currents at the junction resulting in:

$$(v_i - v_\pi)G_s' - v_\pi C_\pi s - v_\pi r_\pi + (v_o - v_\pi)C_\mu s = 0 \quad [16-52]$$

At node v_o , the node equation is:

$$(v_\pi - v_o)C_\mu s - v_o G_L - g_m v_\pi = 0 \quad [16-53]$$

Rewriting and collecting terms results in:

$$\begin{aligned} -v_\pi [G'_s + (C_\pi + C_\mu)s + r_\pi] + v_o C_\mu s &= -v_i G'_s \\ v_\pi [C_\mu s - g_m] - v_o [G_L + C_\mu s] &= 0 \end{aligned} \quad [16-54]$$

This can be put into standard matrix form:

$$\begin{bmatrix} -[G'_s + (C_\pi + C_\mu)s + r_\pi] & C_\mu s \\ C_\mu s - g_m & -[G_L + C_\mu s] \end{bmatrix} \begin{bmatrix} v_\pi \\ v_o \end{bmatrix} = \begin{bmatrix} -v_i G'_s \\ 0 \end{bmatrix} \quad [16-55]$$

Using Cramer's rule, we find the output voltage to be:

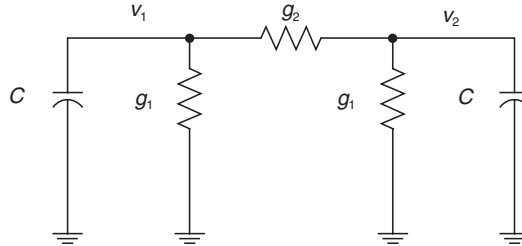
$$v_o = \frac{\det \begin{bmatrix} -[G'_s + (C_\pi + C_\mu)s + r_\pi] & -v_i G'_s \\ C_\mu s - g_m & 0 \end{bmatrix}}{\det \begin{bmatrix} -[G'_s + (C_\pi + C_\mu)s + r_\pi] & C_\mu s \\ C_\mu s - g_m & -[G_L + C_\mu s] \end{bmatrix}} \quad [16-56]$$

This result was summarized in earlier chapters where we found the transfer function for this amplifier.

Finding Oscillation Modes

Consider the circuit in **Figure 16-22**. For convenience, conductances are used instead of resistance, where $g = 1/R$.

Figure 16-22: Symmetric RC circuit.



We can find the natural frequencies of this system by inspection, by considering the cases of “even” and “odd” initial conditions for v_1 and v_2 . For initial conditions when $v_1 = v_2$, there is no current in g_2 and hence we have two independent half circuits, each with natural frequency:

$$\omega_{o,mode1} = \frac{g_1}{C} \quad [16-57]$$

For odd initial conditions with $v_1 = -v_2$, the center point of g_2 is at ground potential, and hence we can ground this point. The resulting natural frequency is then:

$$\omega_{o,mode2} = \frac{g_1 + 2g_2}{C} \quad [16-58]$$

Standard state matrix form for a system of differential equations⁹ is:

$$\{\dot{x}\} = [A]\{x\} \quad [16-59]$$

where $\{x\}$ is the state vector and the *eigenvalues* of the $n \times n$ system matrix $[A]$ are the roots of the characteristic equation (hence, the poles of the system response). The mode shapes are the *eigenvectors* of $[A]$.

In the s -plane, the matrix equations for an RC circuit are:

$$[sC + G]\{v\} = \{i\} \quad [16-60]$$

where $\{i\}$ is the current excitation vector. Natural frequencies (poles) are found by solving the homogenous case, where:

$$[sC + G]\{v\} = \{0\} \quad [16-61]$$

This can be rewritten as:

$$[sC]\{v\} = -[G]\{v\} \quad [16-62]$$

which we recognize as the same state-space formulation! Therefore, the natural frequencies are found by evaluating the determinant of the admittance matrix, or:

$$\det[sC + G] = 0 \quad [16-63]$$

Equivalently,

$$\det[sI - A] = 0 \quad [16-64]$$

where I is the identity matrix and the system matrix $A = G/C$.

Example 16.4: Finding oscillation modes using MATLAB

Next, we'll set up some matrices using the circuit of **Figure 16-22** and solve them using MATLAB. Using KCL, we can derive the state equations for this circuit. For instance, at the v_1 node, KCL gives:

$$-C \frac{dv_1}{dt} - g_1 v_1 + (v_2 - v_1)g_2 = 0 \quad [16-65]$$

Similarly, at the v_2 node:

$$-C \frac{dv_2}{dt} - g_1 v_2 + (v_1 - v_2)g_2 = 0 \quad [16-66]$$

For this system, suitable state variables are v_1 and v_2 , and the state vector is:

$$\{x\} = \begin{Bmatrix} v_1 \\ v_2 \end{Bmatrix} \quad [16-67]$$

In state space form, the equations for the above network is:

$$\begin{Bmatrix} \dot{v}_1 \\ \dot{v}_2 \end{Bmatrix} = - \begin{bmatrix} \frac{g_1 + g_2}{C} & \frac{-g_2}{C} \\ \frac{-g_2}{C} & \frac{g_1 + g_2}{C} \end{bmatrix} \begin{Bmatrix} v_1 \\ v_2 \end{Bmatrix} \quad [16-68]$$

⁹ Remember the notation $\{\dot{x}\}$ being the time derivative of the $\{x\}$ matrix.

Chapter 16

Following is a MATLAB script for finding the natural frequencies of this system, assuming that $g_1 = 1\Omega^{-1}$, $g_2 = 10\Omega^{-1}$ and $C = 1$ farad:

```
function RC

g1=1; g2=10;
% Conductance
C=1;

% Capacitors
% Form A matrix
A=-[(g1+g2)/C -g2/C;
    -g2/C (g1+g2)/C];

% Find natural frequencies and mode shapes
[Modeshapes, NatFreqs] = eig(A);

% Normalize modeshapes
N=size(Modeshapes,1);
for i=1:N
    Modeshapes(:,i)=Modeshapes(:,i)/max(abs(Modeshapes(:,i)));
end

% Display modeshapes
Natural_Frequencies=NatFreqs
Mode_Shapes=Modeshapes
```

And here's the result of running this MATLAB script, where we find the natural frequencies and mode shapes. The natural frequencies are the diagonals of the natural frequency matrix. The mode shapes are the columns of the eigenvector matrix. For instance, for $\omega_o = -1$ radian/second, the mode shape is $[1 \ 1]$, corresponding to the case when initial conditions are set as $v_1 = v_2$. For $\omega_o = -20$ radian/second, the mode shape is $[1 \ -1]$, corresponding to the case when $v_1 = -v_2$.

```
>> RC

Natural_Frequencies =

    -21         0
         0      -1

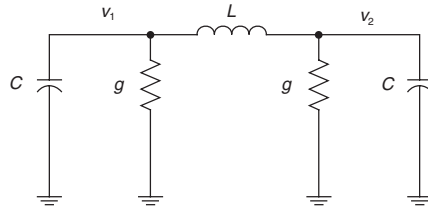
Mode_Shapes =

         1         1
        -1         1
```

Note that the natural frequencies are the same as calculated by the “inspection” method.

We'll repeat the MATLAB formulation using the symmetric LRC circuit of **Figure 16-23**. This circuit has one real and two imaginary poles, as we'll see.

Figure 16-23: Symmetric LRC circuit.



Omitting the details of the node equations, we find the state matrix to be:

$$\begin{Bmatrix} \dot{v}_1 \\ \dot{i} \\ \dot{v}_2 \end{Bmatrix} = - \begin{bmatrix} g/C & 1/C & 0 \\ -1/L & 0 & 1/L \\ 0 & -1/C & g/C \end{bmatrix} \begin{Bmatrix} v_1 \\ i \\ v_2 \end{Bmatrix} \quad [16-69]$$

A MATLAB script follows for solving these simultaneous equations.

```
function LRC

g=1;                % Conductance
C=1;                % Capacitors
L=1;                % inductance

A=-[ g/C 1/C 0
     -1/L 0 1/L
     0 -1/C g/C];    % Form A matrix

% Find natural frequencies and mode shapes
[Modeshapes, NatFreqs] = eig(A);

% Normalize modeshapes
N=size(Modeshapes,1);
for i=1:N
    Modeshapes(:,i)=Modeshapes(:,i)/max(abs(Modeshapes(:,i)));
end

% Display modeshapes
Natural_Frequencies=NatFreqs
Mode_Shapes=Modeshapes

>> lrc

Natural_Frequencies =
    -1.0000         0         0
         0    -0.5000 + 1.3229i         0
         0         0    -0.5000 - 1.3229i
```


Chapter 16

Mode_Shapes =

```

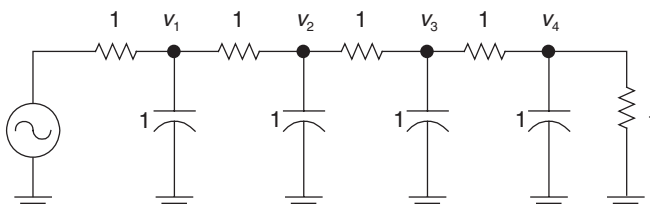
1.0000          0.7071          0.7071
0.0000        -0.3536 - 0.9354i    -0.3536 + 0.9354i
1.0000        -0.7071        -0.7071

```

Note the above natural frequencies and mode shapes. The first natural frequency ($\omega = -1$) has a mode shape of $[1 \ 0 \ 1]$, meaning that $v_1 = v_2$ and that $i = 0$. The second and third natural frequencies are $\omega_{2,3} = -0.5 \pm 1.3229j$, corresponding to a pair of complex poles (and oscillatory behavior for this mode). The mode shapes are also complex, meaning that v_1 , v_2 and i are out of phase for this mode.

Let's next try this method on an RC ladder (**Figure 16-24**).

Figure 16-24:
RC Ladder.



In order to find the natural frequencies and mode shapes of the RC ladder, ground the input and form the G and C matrices. In state-space form this results in:

$$\begin{Bmatrix} \dot{v}_1 \\ \dot{v}_2 \\ \dot{v}_3 \\ \dot{v}_4 \end{Bmatrix} = -\frac{G}{C} \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \begin{Bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{Bmatrix} \quad [16-70]$$

A MATLAB solution is as follows:

```
>> rcladder
```

A =

```

-2    1    0    0
 1   -2    1    0
 0    1   -2    1
 0    0    1   -2

```

Natural_Frequencies =

```

-1.3820    0    0    0
 0   -0.3820    0    0
 0    0   -2.6180    0
 0    0    0   -3.6180

```

Mode_Shapes =

1.0000	-0.6180	-1.0000	-0.6180
0.6180	-1.0000	0.6180	1.0000
-0.6180	-1.0000	0.6180	-1.0000
-1.0000	-0.6180	-1.0000	0.6180

From this analysis, the lowest frequency pole is at $\omega = -0.38$ radians/second. This pole will dominate the input-output response.

Some Comments on Scaling Laws in Nature

Scaling laws and dimensional analysis are very valuable tools for determining how structures, circuits and processes scale. The rationale for using scaling laws is as follows: building a full-scale prototype is often impractical and/or dangerous. An alternate strategy is to build a smaller-scale model and use scaling laws to determine how the full-scale system will behave. This method of scale-modeling is often used in:

- Aerodynamics (wind tunnel tests)
- Hydrodynamics
- Magnetic (i.e., Maglev)
- Rocketry
- Power electronics

Many interesting scaling laws can be inferred by studying simple examples in engineering and nature.

Table 16-2: List of symbols.

B	Magnetic flux density (Tesla)
C_d	Aerodynamic/hydrodynamic drag coefficient
E_k	Kinetic energy = $\frac{1}{2} Mv^2$ (joules)
f_d	Drag force (Newtons)
l	Length scale
P	Power (joules, or watts/sec)
v	Velocity (m/sec)
σ	Electrical conductivity ($1/\Omega\cdot\text{m}$)
ϵ	Dielectric permittivity (farads/meter)
μ	Magnetic permeability (henry/meter)
μ_o	Magnetic permeability of free space ($4\pi \times 10^{-7}$ henry/meter)

Geometric Scaling

By elementary mathematics, in similar figures the surface area scales as the square of the linear dimension and the mass scales as the cube of the linear dimension.¹⁰ In a sphere the surface area is $4\pi r^2$ while the volume is $4/3\pi r^3$ and hence the ratio of volume to surface is $1/3r$. In other words, the ratio of volume to surface area scales as the length scale l .

Fish/ship speed (Froude's law)

If every length scale l is increased by the same factor, how does the speed of a ship (or a fish for that matter) scale? Well, the propulsion power that an engine can provide is proportional to the mass of the engine, which scales as l^3 , or:

$$P_p = k_1 l^3 \quad [16-71]$$

The drag force on the hull is proportional to velocity. The proportionality constant is the drag constant C_d , which is proportional to hull area.

$$f_D = C_d v = k_2 l^2 v \quad [16-72]$$

In steady state, the propulsion power balances the power due to the drag force on the hull, with power equal to force multiplied by velocity. Hence, the drag power is:

$$P_D = f_D v = k_2 l^2 v^2 \quad [16-73]$$

In steady state the propulsion power and the drag power are equal, resulting in:

$$k_1 l^3 = k_2 l^2 v^2 \quad [16-74]$$

Manipulating this expression results in:

$$v^2 = \frac{k_1}{k_2} l \Rightarrow v \propto \sqrt{l} \quad [16-75]$$

This result, known as Froude's law, shows that if everything else is equal the maximum velocity of a ship scales as the square root of the length scale. For example, if a 100-foot long ferry can travel at 10 mph, a scaled-up ship that is 1000 feet long will have a maximum speed of approximately 32 mph.

Fruit

How do the sizes of plants scale? For instance, the stalk of a piece of fruit can withstand a maximum stress (or force per unit area) in Newtons/m². The force exerted on the stalk scales as the mass of the fruit, or:

$$f \propto l^3 \quad [16-76]$$

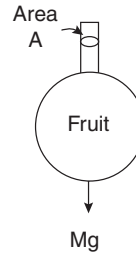
The area of the stalk scales as l^2 . Therefore, the force/area scales as l . Results from this scaling law are:

- Enormous structures, if scaled up directly, will collapse under their own weight.
- Old trees have huge bases.

¹⁰ This example, as well as some of the examples to follow, are derived from Darcy W. Thompson's, *On Growth and Form*.

- The Eiffel tower is tapered.
- Large fruit lie on the ground.

Figure 16-25: Fruit with stalk.



Bending moments

By elementary mechanics, the bending of a beam under its own weight scales as the length squared, if the cross-sectional area remains constant. Said another way, a 12-foot 2×4 fixed at one end will have four times the deflection of a 6-foot 2×4 fixed at one end. In nature, to counteract this effect, as the size of an animal increases the limbs tend to become thicker and shorter.

Size and heat in bodies (Bergman's law)

Heat production in a body (i.e., power) is proportional to mass, or l^3 . Heat loss is proportional to surface area, or l^2 . The ratio

$$\frac{\text{Heat loss}}{\text{Heat production}} \propto \frac{1}{l} \quad [16-77]$$

shows that small animals have proportionally higher heat loss than large animals. This is why birds and mice need to eat lots of food in proportion to their body weight. A man eats 1/50 of his body weight per day, while a mouse eats approximately half his body weight in food. Another result is that there aren't any tiny animals in the arctic.

Size and jumping (Borelli's law)

When an animal jumps, the leg muscles impart a force impulse to the animal; this impulse is proportional to muscle mass, or l^3 . Therefore, the velocity is:

$$v \propto \frac{\text{Muscular force impulse}}{\text{Mass moved}} = \text{constant} \quad [16-78]$$

Borelli's law shows why all animals can jump approximately the same height.

Walking speed (Froude's law)

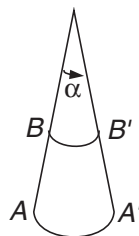
Assuming that everyone's legs travel through the same angle α , a model for determining the scaling law for walking speed is found in **Figure 16-26**. The distance per oscillation for one step is proportional to A/B . Assuming that the legs behave as pendulums, the time per oscillation scales as:

$$t \propto \sqrt{\frac{A}{B}} \quad [16-79]$$

The walking speed v is found by:

$$v \propto \frac{\text{distance}}{\text{time}} \propto \sqrt{\frac{A}{B}} \quad [16-80]$$

Figure 16-26: Model for determining walking speed.



This again is Froude's law.

Capacitors

How does the RC time constant of a capacitor scale with size? A model for a parallel-plate capacitor is shown in **Figure 16-27a**. Parallel plates are spaced d apart, with a material inside with dielectric permittivity ε and electrical conductivity σ . An electrical model for this capacitor is shown in **Figure 16-27b**. The capacitance is given by:

$$C = \frac{\varepsilon A}{d} \quad [16-81]$$

where A is the area of the plates. The resistance is:

$$R = \frac{d}{\sigma A} \quad [16-82]$$

Therefore, the RC time constant is

$$RC = \frac{\varepsilon}{\sigma} \quad [16-83]$$

independent of size.

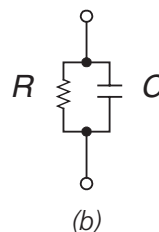
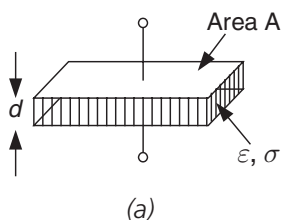


Figure 16-27: Capacitor model.

Inductors

A model for a single-loop inductor is shown in **Figure 16-28**.

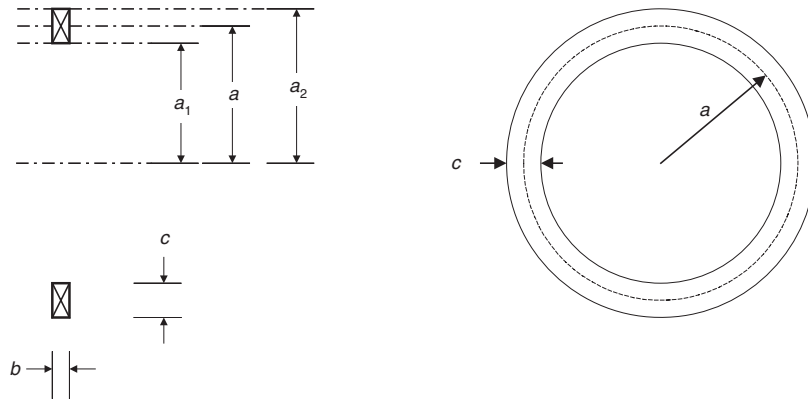


Figure 16-28: Inductor model.

The inductor has mean radius a , axial thickness b and radial width c . For a thin hoop with $c \ll a$ the resistance of the loop is:

$$R \approx \frac{2\pi a}{bc\sigma} \quad [16-84]$$

The inductance of the loop is found by using the inductance reference by Frederick Grover, *Inductance Calculations*:

$$L \approx kaPF \quad [16-85]$$

with

$k = \text{constant}$

$P = \text{function of } c/2a$

$F = \text{function of } b/a$

The time constant of an inductor is L/R , which is found by:

$$\frac{L}{R} \approx \frac{kPFbc\sigma}{2\pi} \propto l^2 \quad [16-86]$$

An interesting result pops out here: magnetic scaling laws show that large magnetic elements are more efficient in energy conversion than smaller ones. For this disk coil geometry, this effect can be quantified by considering the ratio of inductance to resistance. The inductance L is approximately proportional to a as shown above. The resistance of the coil is proportional to a/bc , the ratio of current path length to coil cross-sectional area. Therefore, the ratio of inductance to resistance is proportional to bc , or the cross-sectional area of the coil. If all coil lengths are scaled up by the same factor l , this ratio increases by the factor l^2 , or the length squared. This scaling law shows that the efficiency of inductors (and hence, electric motors) improves with increasing size as l^2 .

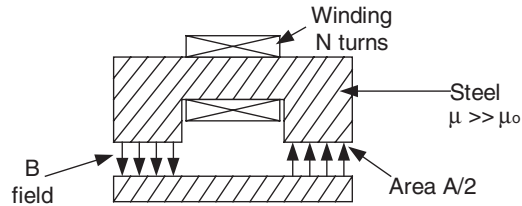
Lift force of electromagnet

The lift force f_l of an electromagnet (**Figure 16-29**) is found by analyzing the magnetic pressure, and is approximately:

$$f_l \approx \frac{B^2 A}{2\mu_o} \quad [16-87]$$

where B is magnetic flux density (Tesla), A is the total area of the two poles of the electromagnet, and μ_o is the magnetic permeability of free space.

Figure 16-29: Electromagnet.



How does this force scale with magnet length scale? The B field is found approximately by:

$$B \approx \frac{\mu_o NI}{2g} \propto \frac{N}{l} \quad [16-88]$$

where g is the airgap. The number of turns scales with the area, or:

$$N \propto l^2 \quad [16-89]$$

The lift force scales as $B^2 A$, or as l^4 , while the weight of the magnet scales as l^3 . The lift/weight ratio therefore scales as l . We might infer from this that a tiny electromagnet cannot lift its own weight.

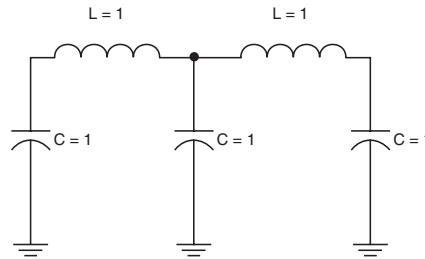
Chapter 16 Problems

Problem 16.1

For the LC circuit in **Figure 16-30**:

- (a) By analogy, find the natural frequencies of the LC circuit.
- (b) Simulate your results using SPICE. Excite each mode independently by proper setting of inductor current initial conditions.

Figure 16-30: Circuit for Problem 16.1.



Problem 16.2

Using Cramer's rule, solve the following simultaneous linear equations:

$$x + y = -1$$

$$2x - 3y = 13$$

Problem 16.3

Marc is 6 feet tall and Lisa is 5 feet tall.¹¹ Marc walks a mile in 20 minutes at his most comfortable walking speed. At her most comfortable walking speed, approximately how long does it take Lisa to walk a mile? Roughly how long does it take their 2-year-old daughter Sophie to walk a mile?

Problem 16.4

Find the thermal resistance through an aluminum plate with dimensions T (in the direction of heat flow) = 1 millimeter, and cross-sectional area $A = 100 \text{ mm}^2$. Assume that the thermal conductivity of aluminum $k = 240 \text{ watts/(m}^\circ\text{C)}$.

¹¹ Names changed to protect the innocent.

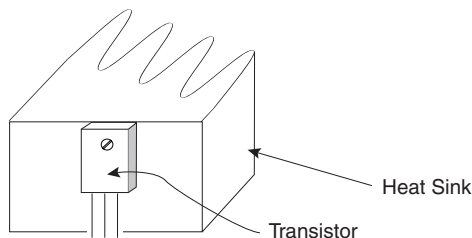
Problem 16.5

A 2N3906 transistor is in a circuit in which the transistor dissipates 100 mW of energy. The ambient temperature is 40°C. What is the junction temperature of the device? Assume that the thermal resistance, junction to ambient, is 200°C/watt. Draw the thermal circuit.

Problem 16.6

- Draw an appropriate *static* thermal model for the transistor mounted to a heat sink (**Figure 16-31**). Include junction temperature, case temperature, and heat sink temperature.
- Assume that total loss in the transistor is 10 watts. Assuming the following parameters: $T_A = 50^\circ\text{C}$; $R_{THjc} = 1^\circ\text{C/W}$; $R_{THcs} = 1^\circ\text{C/W}$; Calculate the maximum thermal resistance from heat sink to ambient, R_{THsa} , to keep the junction temperature of the transistor below 150°C.

Figure 16-31: Transistor mounted to a heat sink for Problem 16.6.

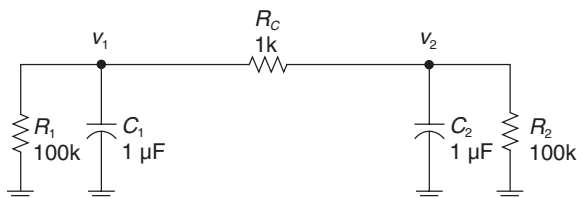


Problem 16.7

Using the symmetric RC circuit of **Figure 16-32**:

- Assume that at time $t = 0$ that the circuit has initial conditions $v_1 = v_2 = 1\text{V}$. Sketch the responses of $v_1(t)$ and $v_2(t)$ for $t > 0$.
- Assume that at $t = 0$ that $v_1 = 1$ and $v_2 = -1\text{V}$. Sketch the responses of $v_1(t)$ and $v_2(t)$ for $t > 0$.

Figure 16-32: Circuit for Problem 16.7.



Problem 16.8

The engine block in an automobile has the following parameters:

- Mass: 100 kg of aluminum
 - Power dissipated when idling: 5000W
 - Aluminum heat capacity $c_p = 900\text{ J/kg}\cdot^\circ\text{C}$
- When you first turn the engine on, the engine is at ambient temperature ($T_A = 25^\circ\text{C}$) and there is no water circulating through the engine (since the engine is cold). With

no water circulation, the thermal resistance from engine-to-ambient is 0.1°C per watt. Generate a lumped thermal model of the system, assuming that the temperature of the motor is uniform throughout. Note that thermal capacitance is Mc_p (joules/ $^{\circ}\text{C}$).

- (b) The engine is “warmed up” when the temperature of the engine reaches 100°C . How long does the engine take to warm up?
- (c) During running, assume that the temperature of the engine reaches a final value of 125°C . Plot the temperature of the engine after it is shut off.

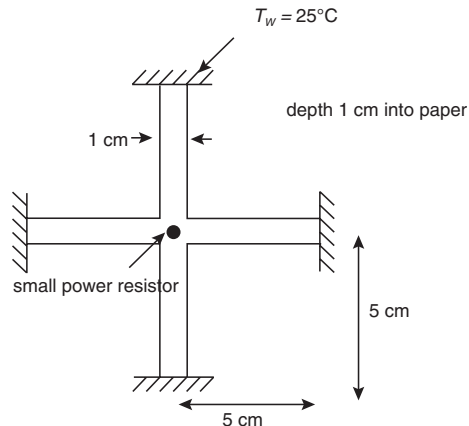
Problem 16.9

This problem concerns a power resistor mounted in a cross-shaped block of copper, as shown in **Figure 16-33**. The area of each member is $1\text{ cm} \times 1\text{ cm}$ as shown, with the length of each member being 5 cm long. The cross-sectional area of the resistor is small compared to the area of the copper (hence you can consider the resistor to be a point source of power). The four members mount to an outside wall that has a constant temperature of 25°C . A layer of thermal grease with average thickness 0.02 mm is used to thermally couple the edge of the copper bars to the wall. Since this layer is thin, you may neglect the thermal capacitance of the thermal grease layer.

- (a) Generate a lumped thermal model of the system.
- (b) If you energize the resistor at $t = 0$ with 100 W of power, what is the final resistor temperature as $t \rightarrow \infty$? (Assume that the thermal joint between the resistor and the copper members is perfect.)
- (c) In a separate experiment, before $t = 0$, the wall, the copper and the resistor are at a uniform temperature and there is no current flow in the power resistor. Assume that the power resistor is energized and dissipates 100 W , beginning at $t = 0$. Sketch the resistor temperature for all time, showing breakpoints, time constants, etc. given your thermal model and the simplifying assumptions made.

(Hint: parts (b) and (c) can be made significantly simpler by exploiting symmetry, breaking up the large problem into smaller chunks, using intuitive reasoning as to the power flow paths, etc.).

Figure 16-33: Thermal system for Problem 16.9.



References

- Fogel, M., Editor, *The Automatic Control Systems/Robotics Problem Solver*, Research and Education Association, Piscataway NJ, 1990.
- General Electric Corp., *SCR Handbook, 6th Edition. Good general-purpose manual with lots of practical advice on heat sink design, and mounting and cooling techniques for power devices.*
- Gilbert, B., “A New Wide-Band Amplifier Technique,” *IEEE Journal of Solid State Circuits*, vol. SC-3, no. 4, December 1968, pp. 353–365.
- , “A Precise Four-Quadrant Multiplier with Subnanosecond Response,” *IEEE Journal of Solid State Circuits*, vol. SC-3, no. 4, December 1968, pp. 365–373.
- , “A DC-500 MHz Amplifier Multiplier Principle,” *1968 ISSCC Digest of Technical Papers*, pp. 114–115.
- , “Translinear Circuits: A Proposed Classification,” *Electronics Letters*, vol. , no. 1, pp. 14–16, 1975.
- , “Current-Mode Circuits from a Translinear Viewpoint: A Tutorial,” in C. Toumazou, F. J. Lidgey, and D. G. Haigh, eds., *Analogue IC Design: The Current-Mode Approach*, London: Peter Peregrinus, pp. 11–91, 1990.
- , “Translinear Circuits—25 Years On, Part I: The Foundations,” *Electronic Engineering*, vol. 65, no. 800, pp. 21–24, 1993.
- , “Translinear Circuits: An Historical Review,” *Analog Integrated Circuits and Signal Processing*, vol. 9, no. 2, pp. 95–118, 1996.
- Grover, F., *Inductance Calculations*, originally published by D. Van Nostrand, 1946, reprinted by Dover, 2004.
- Jezierski, E., “On Electrical Analogues of Mechatronic Systems,” *Proceedings of the 2001 Second International Workshop on Motion and Control*, October 18–20, 2001, pp. 181–188.
- Kassakian, J., Schlecht, M., and Verghese, G., *Principles of Power Electronics*, Prentice-Hall, 1991. *A well-regarded reference for all aspects of power electronics design, including magnetics.*
- Klein, Richard E., “Teaching Linear Systems Theory Using Cramer’s Rule,” *IEEE Transactions on Education*, vol. 33, no. 3, August 1990, pp. 258–267.
- Langlois, P. J., “Graphical Analysis of Delay Line Waveforms: A Tutorial,” *IEEE Transactions on Education*, vol. 38, no. 1, February 1995, pp. 27–32.
- Mathworks, Inc., *MATLAB Reference Guide*, 1992.
- Middlebrook, R. D., “Null Double Injection and the Extra Element Theorem,” *IEEE Transactions on Education*, vol. 32, no. 3, August 1989, pp. 167–180.
- , “The Two Extra Element Theorem,” *Proceedings of the IEEE Frontiers in Education Twenty-First Annual Conference*, Purdue University, September 21–24, 1991, pp. 702–708.
- , “Low-Entropy Expressions: The Key to Design-Oriented Analysis,” *Proceedings of the IEEE Frontiers in Education Twenty-First Annual Conference*, Purdue University, September 21–24, 1991, pp. 399–403.

- Rao, Singiresu S., *Mechanical Vibrations*, 3d edition, Addison-Wesley, 1995.
- Rothkopf, Elliot, “Teaching for Understanding—Analogies for Learning in Electrical Engineering,” *1995 IEEE Frontiers in Education Conference*, session 2b4, pp. 2b4.9–2b4.13.
- Seevinck, E., and Wiegerink, R. J., “Generalized translinear circuit principle,” *IEEE Journal of Solid-State Circuits*, Volume 26, no. 8, August 1991, pp. 1098–1102.
- Shen, Liang Chi, and Kong, Jin Au, *Applied Electromagnetism*, Brooks/Cole, 1983.
- Teoh, Chin Soon, and Davis, Lionel E., “A Coupled Pendula System as an Analogy to Coupled Transmission Lines,” *IEEE Transactions on Education*, vol. 39, no. 4, November 1996, pp. 548–557.
- Thompson, D’Arcy W., *On Growth and Form*, Dover Publications (reprinted), 1992.
- Thornton, Richard D., *Electronic Circuits Modeling, Analysis, Simulation; Intuition and Design*, MIT course notes, January 1993.
- Vorperian, Vatche, “Improved Circuit-Analysis Techniques Require Minimum Algebra,” *EDN*, August 3, 1995, pp. 125–134.
- Zahn, Markus, *Electromagnetic Field Theory: A Problem Solving Approach*, reprinted by Krieger, 1987.

Index

A

Acceptors 53–55

Amplifiers

Amplifier, cascode 146–147, 191, 251, 355

Amplifier, common-base 141, 191–194

Amplifier, common-emitter 101–106, 118, 122, 132–135, 139, 142, 146, 157–158, 165, 186–188, 202, 207–208, 212, 228–229, 457

Amplifier, common-emitter with emitter degeneration 165–167, 203, 358

Amplifier, current 194

Amplifier, differential 112–118, 139–140, 167–168, 248, 348, 355, 361–362

Amplifier, high-gain 225, 227, 229, 231, 233–234, 236

Amplifier, logarithmic 6–7

Amplifier, MOS 241, 248, 255–256, 258, 261, 264–265

Amplifier, peaking 118, 190–191

Amplifier, photodiode 327–330

Amplifier, push-pull 351, 358

Amplifier, video 203, 344, 379

Analogies

Mechanical circuits 117, 444–448

thermal circuits 440–448

B

Biasing 97–125

Bode plot 36–37, 142, 175, 185, 314, 316–318, 320, 323–324, 326, 332, 375, 390

Bardeen, John 3, 43, 75

Base spreading resistance 86, 92–93, 101, 143, 187, 190, 236, 290, 373

Base-width modulation 207–211, 246, 373

Beta 78, 84, 88, 90, 220–221, 358

Biasing, transistor 97

Black, Harold 306–308

Bootstrapping 165, 177–183

Brattain, Walter 3, 75

C

Capacitance

Capacitance, diffusion 63, 65–66, 85

Capacitance, junction 63–64, 85–86, 91–93, 274–276, 278–279, 282, 290

Capacitors

Parasitic elements 421, 424, 436

ESR 422, 436

Cascaded systems 37

Cascode 146, 150, 153, 191, 214–215, 222, 251, 254, 262, 355

Charge control model 269–303

Characteristic impedance 25, 34–35, 418, 448, 452

Common mode 115–118

Common-source amplifier 249–251, 255

Conductive heat transfer 440

Conductor 43–45

Cramer's rule 103, 111, 197, 455–458

Crossover distortion 348–349, 351

Crossover frequency 314, 316–317, 321–322, 325–326, 328, 330, 332

Current density 46, 61–62

Current feedback 369–381
Current feedback op-amps 5, 194, 369–381
 Absence of slew rate limiting 375
 Basic current feedback op-amp 369, 372
 LM6181 379–380
 Current-feedback op-amp limitations 5, 369, 381
Current mirror 218–237, 352–354, 358, 372–373, 376, 379
 Basic current mirror 218–219
 Cascode current mirror 222
 Current mirror with “beta helper” 220
 Current mirror with emitter degeneration 219
 MOS current mirrors 252–254
 Wilson current mirror 221, 252–254
 Widlar current mirror 222–224
Current repeater 218
Current source 15, 72, 108, 110, 114–117, 133, 136, 144, 161, 164, 171, 174, 208, 210, 212–215, 219, 224–225, 228, 230–231, 250–252, 258, 278, 327, 331–334, 352, 355, 358, 376, 379, 425, 429–430, 442, 444
 Current source, incremental output resistance 115, 208, 212–213, 224–225, 231, 250, 253, 348
 MOSFET current source 331–334, 340
Current switching 194–196, 292, 431–432
Current gain 75–76, 78, 84, 86–91, 97–98, 100–101, 109, 133, 139, 146, 165, 169, 187, 211–212, 216, 219–220, 228, 295–297, 372

D

Damping ratio 25–30, 315, 321, 324, 337
Datasheet
 Datasheet, diode 67–69
 Datasheet, finding parameters on 211–212
 Datasheet h-parameters 211–212
 Datasheet, transistor 84, 88–93, 211–212
Depletion region 54–63, 65–67, 71, 89, 210, 246, 282, 287
Depletion capacitance 63–64, 85, 91, 247, 274, 287

Desensitivity in feedback systems 309
Differential mode 112, 114–116, 118, 139–140, 167, 194, 348, 362
Diffusion 48, 52, 54–57, 59–63, 65–67, 77, 85, 270
 Diffusion constants 50–51, 62
Diode
 Ideal diode equation 62–63
 In thermal equilibrium 54–57, 77
 Forward-biased 57–58, 60, 77–79, 81, 282
 Reverse-biased 61, 66, 77–80, 271
 Schottky 70
Disturbance rejection in feedback systems 309
Donors 53–55, 59
Doping 43, 48, 52–56, 59, 276
Drift 48–52, 54, 56–57, 77, 360–361

E

Early effect 224
Electrical conductivity 43–44, 242, 421, 440, 463, 466
Electron 44–54, 59–61, 65–66, 77
Electron current 60–61
Emitter degeneration 165–168, 219–220, 224, 358
Emitter follower 107–112, 117, 136–139, 150, 153, 215–217, 228, 230–233, 248, 263, 265, 348, 355
 Input and output impedance 107–110, 168–176, 215
Energy method 34–35, 41

F

Feedback systems 305–344
 Crossover frequency 314, 316–317, 321–322, 325–326, 328, 330, 372
 Desensitivity 309
 History of feedback control 305
 Loop transmission 309, 314, 316–318, 320–322, 325–326, 328, 332–333
 Phase margin and gain margin 314–315, 317, 321–323, 325–326, 332–333
 Routh stability criterion 311, 319

Stability 310–311, 314, 319, 328, 330

Fick's law 49

Filters

All-pass 406–407

Bessel 395–401

“Brick wall” filter 404–405

Butterworth 387–389

Chebyshev 390–395

Comparison of different filter types 398–399

Elliptic 404–405, 408

Group delay 385–390, 394–396, 398–399, 404, 406–408, 410–413

Implementation 400, 404

Ladder 400, 403, 408–413

Passband 386–387, 390–396, 398, 402, 404, 408, 410–411

Sallen-Key 387, 404, 406, 408

Stopband 386, 398, 404, 410

Transition band 388, 390, 395, 405

Full-wave rectifier 20–21

Fusing current 426–427

G

Gain-bandwidth product 5, 86–87, 139, 297, 332, 347, 369–371, 375, 378, 411

Gain margin 314–315, 317, 326

Generation 48, 52, 54, 78

Group delay 18–19, 385–390, 394–396, 398–399, 404, 406–408, 410–413

Grover, Frederick W., 428, 467

H

Hole 45–62, 65–67, 77–79, 270

Hole current 46, 60–61

Hybrid- π model 207, 209–212, 215, 222, 269, 274

Extended hybrid- π model 210–212, 215, 222

Hybrid- π model, limitations 93

I

Incremental model 82, 84–87, 165, 169, 171, 173–174, 207–208, 215, 226, 252–253

Inductors 2, 7, 14–15, 23, 34, 400, 408, 424, 444, 447, 452, 467

Nonideal 424

Insulator 43–45, 441

Intrinsic carrier 47, 59, 71

Intrinsic carrier concentration 47, 59, 71

K

Kilby, J., and invention of the integrated circuit 4

L

Lag network 316

Laplace transform 9, 13

Laser diode driver 428

Lead network 317

Lilienfeld, Julius 2–3, 241

Load line 43, 71–72, 80, 97–98

Loop transmission 309, 314, 316–318, 320–322, 325–326, 328, 332–333

Low-pass filters 385–415

M

Maglev 334, 336–337, 341, 463

Mass-spring system 23–24, 273

MATLAB xii, 33, 104, 132, 321, 326, 338, 340–341, 392, 459–462

Miller effect 105–107, 145–147, 177, 194, 251, 262, 372

Minority carriers 51, 59–60, 270

Minority carrier concentration 59, 61, 65, 85, 210, 270, 276

MOSFET

“Back-gate” effect 247–251, 255

Basic MOS devices 241–267

Early history 241

V/I curve 243–246

Small-signal model 246–247, 255–256, 258–259, 262–263

N

Natural frequency 25–28, 30, 444, 446–447, 458, 460
Node equations 455–457
Nonsaturating current switch 290

O

Open-circuit time constants 127–155
 Worst case calculation 161–165
Operational amplifier 4
 Bias current 362
 Capacitive loading 363–365
 Common-mode rejection ratio (CMRR) 117
 Current-feedback op-amps 369–381
 Gain-bandwidth product 369–371, 375, 378
 Output resistance 363
 Slew rate 362–363, 371, 375
 Voltage offset 359–360
 Voltage offset drift 360–361
 741 op-amp 358
Oscillation modes 458–463

P

PC boards, (see printed-circuit boards)
Phase margin 314–315, 321–326, 365
Pole/zero plot 36–37
Pole splitting 196–197, 199–200, 348, 358, 362
Printed-circuit boards 425–427
 Design case study 428–435
 Ground planes 426
 Layout issues 425–427
 Power supply bypassing 425
 Trace widths 426

Q

Q (see “quality factor”)
Quality factor 27–28

R

Recombination 48, 52, 54, 89, 280–281, 286
Resonance 27–28, 33–35, 430
Resistors 417
 Parasitic elements 417–419
 Surface-mount resistors 419–420
 Temperature coefficients 420–421
Resistive ladder 451
Reverse breakdown 66–68
Reverse recovery 65–70
Risetime 16–17, 22, 30, 37, 277, 287, 289, 320–322, 325, 334, 378, 430–435
Routh stability criterion 311, 319

S

Saturation 272–273, 289
Saturation delay 286
Scaling laws 463, 467
Second-order systems 23–32, 129, 198, 310–311, 315
Semiconductor 43–54
Shockley, William 3, 43, 45, 51–52, 62, 75–76
Short-circuit time constants 142, 184–190, 228
Silicon lattice 45–47, 55
Source follower 248–249, 257–260, 264–265, 331
 Step response 19–30

T

Thermal circuits 439–444
Thermal equilibrium 54–57, 77
Transfer function 36–37, 87, 104, 111, 119, 127–129, 135, 138, 184–186, 190, 199–200, 221, 250, 273, 308–310, 312–313, 316–317, 319–337, 363, 367, 369, 387–388, 390, 392–393, 397–398, 407, 430, 455, 457–458
Transistor
 Transistor, biasing 97–101
 Transistor, cutoff region 79, 282–284

-
- Transistor, forward-active 270–272, 278, 281, 283–284, 287
 - Transistor inverter 277–289
 - Transistor, point contact 75–76
 - Reverse-active 271–272
 - Transistor, saturation region 272–273
 - Transistor datasheet 211–212
 - Transistor switching 269–303
 - Emitter switching 292–293
 - Inverter 277–289
 - Nonsaturating current switch 290–291
 - Translinear principle 449–450
 - Transmission lines 451–455
- W**
- Widlar, Robert 5–6
 - Widely-spaced pole approximation 185, 273
- Z**
- Z_o (see “characteristic impedance”)

What's on the CD-ROM?

Included on the CD-ROM are Powerpoint presentations for each chapter, suitable for use by instructors. Also included are selected MATLAB scripts which were used in preparation of the text.

ELSEVIER SCIENCE CD-ROM LICENSE AGREEMENT

PLEASE READ THE FOLLOWING AGREEMENT CAREFULLY BEFORE USING THIS CD-ROM PRODUCT. THIS CD-ROM PRODUCT IS LICENSED UNDER THE TERMS CONTAINED IN THIS CD-ROM LICENSE AGREEMENT ("Agreement"). BY USING THIS CD-ROM PRODUCT, YOU, AN INDIVIDUAL OR ENTITY INCLUDING EMPLOYEES, AGENTS AND REPRESENTATIVES ("You" or "Your"), ACKNOWLEDGE THAT YOU HAVE READ THIS AGREEMENT, THAT YOU UNDERSTAND IT, AND THAT YOU AGREE TO BE BOUND BY THE TERMS AND CONDITIONS OF THIS AGREEMENT. ELSEVIER SCIENCE INC. ("Elsevier Science") EXPRESSLY DOES NOT AGREE TO LICENSE THIS CD-ROM PRODUCT TO YOU UNLESS YOU ASSENT TO THIS AGREEMENT. IF YOU DO NOT AGREE WITH ANY OF THE FOLLOWING TERMS, YOU MAY, WITHIN THIRTY (30) DAYS AFTER YOUR RECEIPT OF THIS CD-ROM PRODUCT RETURN THE UNUSED CD-ROM PRODUCT AND ALL ACCOMPANYING DOCUMENTATION TO ELSEVIER SCIENCE FOR A FULL REFUND.

DEFINITIONS

As used in this Agreement, these terms shall have the following meanings:

"Proprietary Material" means the valuable and proprietary information content of this CD-ROM Product including all indexes and graphic materials and software used to access, index, search and retrieve the information content from this CD-ROM Product developed or licensed by Elsevier Science and/or its affiliates, suppliers and licensors.

"CD-ROM Product" means the copy of the Proprietary Material and any other material delivered on CD-ROM and any other human-readable or machine-readable materials enclosed with this Agreement, including without limitation documentation relating to the same.

OWNERSHIP

This CD-ROM Product has been supplied by and is proprietary to Elsevier Science and/or its affiliates, suppliers and licensors. The copyright in the CD-ROM Product belongs to Elsevier Science and/or its affiliates, suppliers and licensors and is protected by the national and state copyright, trademark, trade secret and other intellectual property laws of the United States and international treaty provisions, including without limitation the Universal Copyright Convention and the Berne Copyright Convention. You have no ownership rights in this CD-ROM Product. Except as expressly set forth herein, no part of this CD-ROM Product, including without limitation the Proprietary Material, may be modified, copied or distributed in hardcopy or machine-readable form without prior written consent from Elsevier Science. All rights not expressly granted to You herein are expressly reserved. Any other use of this CD-ROM Product by any person or entity is strictly prohibited and a violation of this Agreement.

SCOPE OF RIGHTS LICENSED (PERMITTED USES)

Elsevier Science is granting to You a limited, non-exclusive, non-transferable license to use this CD-ROM Product in accordance with the terms of this Agreement. You may use or provide access to this CD-ROM Product on a single computer or terminal physically located at Your premises and in a secure network or move this CD-ROM Product to and use it on another single computer or terminal at the same location for personal use only, but under no circumstances may You use or provide access to any part or parts of this CD-ROM Product on more than one computer or terminal simultaneously.

You shall not (a) copy, download, or otherwise reproduce the CD-ROM Product in any medium, including, without limitation, online transmissions, local area networks, wide area networks, intranets, extranets and the Internet, or in any way, in whole or in part, except that You may print or download limited portions of the Proprietary Material that are the results of discrete searches; (b) alter, modify, or adapt the CD-ROM Product, including but not limited to decompiling, disassembling, reverse engineering, or creating derivative works, without the prior written approval of Elsevier Science; (c) sell, license or otherwise distribute to third parties the CD-ROM Product or any part or parts thereof; or (d) alter, remove, obscure or obstruct the display of any copyright, trademark or other proprietary notice on or in the CD-ROM Product or on any printout or download of portions of the Proprietary Materials.

RESTRICTIONS ON TRANSFER

This License is personal to You, and neither Your rights hereunder nor the tangible embodiments of this CD-ROM Product, including without limitation the Proprietary Material, may be sold, assigned, transferred or sub-licensed to any other person, including without limitation by operation of law, without the prior written consent of Elsevier Science. Any purported sale, assignment, transfer or sublicense without the prior written consent of Elsevier Science will be void and will automatically terminate the License granted hereunder.

TERM

This Agreement will remain in effect until terminated pursuant to the terms of this Agreement. You may terminate this Agreement at any time by removing from Your system and destroying the CD-ROM Product. Unauthorized copying of the CD-ROM Product, including without limitation, the Proprietary Material and documentation, or otherwise failing to comply with the terms and conditions of this Agreement shall result in automatic termination of this license and will make available to Elsevier Science legal remedies. Upon termination of this Agreement, the license granted herein will terminate and You must immediately destroy the CD-ROM Product and accompanying documentation. All provisions relating to proprietary rights shall survive termination of this Agreement.

LIMITED WARRANTY AND LIMITATION OF LIABILITY

NEITHER ELSEVIER SCIENCE NOR ITS LICENSORS REPRESENT OR WARRANT THAT THE INFORMATION CONTAINED IN THE PROPRIETARY MATERIALS IS COMPLETE OR FREE FROM ERROR, AND NEITHER ASSUMES, AND BOTH EXPRESSLY DISCLAIM, ANY LIABILITY TO ANY PERSON FOR ANY LOSS OR DAMAGE CAUSED BY ERRORS OR OMISSIONS IN THE PROPRIETARY MATERIAL, WHETHER SUCH ERRORS OR OMISSIONS RESULT FROM NEGLIGENCE, ACCIDENT, OR ANY OTHER CAUSE. IN ADDITION, NEITHER ELSEVIER SCIENCE NOR ITS LICENSORS MAKE ANY REPRESENTATIONS OR WARRANTIES, EITHER EXPRESS OR IMPLIED, REGARDING THE PERFORMANCE OF YOUR NETWORK OR COMPUTER SYSTEM WHEN USED IN CONJUNCTION WITH THE CD-ROM PRODUCT.

If this CD-ROM Product is defective, Elsevier Science will replace it at no charge if the defective CD-ROM Product is returned to Elsevier Science within sixty (60) days (or the greatest period allowable by applicable law) from the date of shipment.

Elsevier Science warrants that the software embodied in this CD-ROM Product will perform in substantial compliance with the documentation supplied in this CD-ROM Product. If You report significant defect in performance in writing to Elsevier Science, and Elsevier Science is not able to correct same within sixty (60) days after its receipt of Your notification, You may return this CD-ROM Product, including all copies and documentation, to Elsevier Science and Elsevier Science will refund Your money.

YOU UNDERSTAND THAT, EXCEPT FOR THE 60-DAY LIMITED WARRANTY RECITED ABOVE, ELSEVIER SCIENCE, ITS AFFILIATES, LICENSORS, SUPPLIERS AND AGENTS, MAKE NO WARRANTIES, EXPRESSED OR IMPLIED, WITH RESPECT TO THE CD-ROM PRODUCT, INCLUDING, WITHOUT LIMITATION THE PROPRIETARY MATERIAL, AND SPECIFICALLY DISCLAIM ANY WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

If the information provided on this CD-ROM contains medical or health sciences information, it is intended for professional use within the medical field. Information about medical treatment or drug dosages is intended strictly for professional use, and because of rapid advances in the medical sciences, independent verification of diagnosis and drug dosages should be made.

IN NO EVENT WILL ELSEVIER SCIENCE, ITS AFFILIATES, LICENSORS, SUPPLIERS OR AGENTS, BE LIABLE TO YOU FOR ANY DAMAGES, INCLUDING, WITHOUT LIMITATION, ANY LOST PROFITS, LOST SAVINGS OR OTHER INCIDENTAL OR CONSEQUENTIAL DAMAGES, ARISING OUT OF YOUR USE OR INABILITY TO USE THE CD-ROM PRODUCT REGARDLESS OF WHETHER SUCH DAMAGES ARE FORESEEABLE OR WHETHER SUCH DAMAGES ARE DEEMED TO RESULT FROM THE FAILURE OR INADEQUACY OF ANY EXCLUSIVE OR OTHER REMEDY.

U.S. GOVERNMENT RESTRICTED RIGHTS

The CD-ROM Product and documentation are provided with restricted rights. Use, duplication or disclosure by the U.S. Government is subject to restrictions as set forth in subparagraphs (a) through (d) of the Commercial Computer Restricted Rights clause at FAR 52.22719 or in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.2277013, or at 252.2117015, as applicable. Contractor/Manufacturer is Elsevier Science Inc., 655 Avenue of the Americas, New York, NY 10010-5107 USA.

GOVERNING LAW

This Agreement shall be governed by the laws of the State of New York, USA. In any dispute arising out of this Agreement, you and Elsevier Science each consent to the exclusive personal jurisdiction and venue in the state and federal courts within New York County, New York, USA.